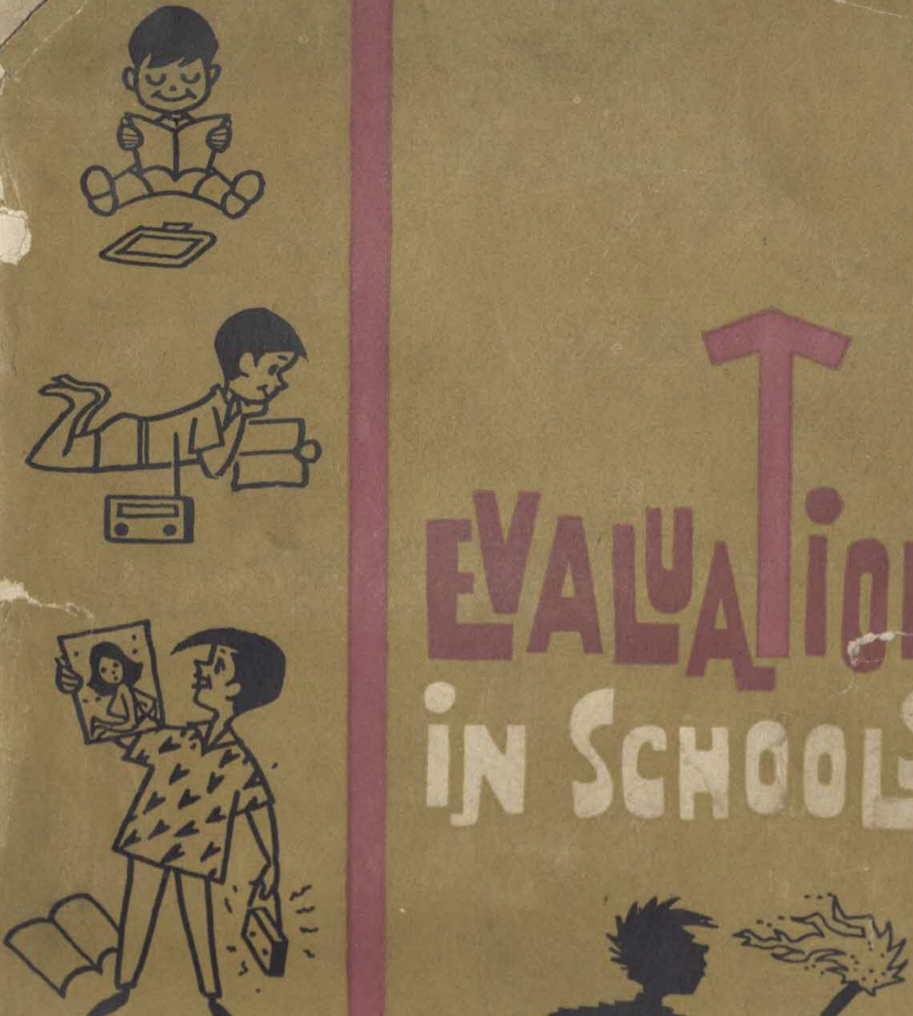
EVALUATION IN SCHOOLS



W. N. DANDEKAR

EVALUATION IN SCHOOLS

by
W. N. DANDEKAR

श्री विद्या प्रकाशन

SHRIVIDYA PRAKASHAN
250 Shaniwar, Ashtabhuja Marg, Poona 30

4727
24.10.89

To,
The Director,
State Council of Educational
Research and Training
25/3, Ballygunge
Circular Road, Calcatta
for the recommendation

— Publisher
19.9.89

EVALUATION IN SCHOOLS

W. N. Dandekar

Formerly Principal S. T. College, Bombay
and
Ex-Director of Education
(Goa, Daman and Div)

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

SHRI VIDYA PRAKASHAN, POONA-30

C.E R.T., West Bengal
24 10 89
ec. No. 4727

371.26
DAN

Copyrights reserved

First Edition : Oct. 1971
Reprinted : 1977
Second Reprint : Nov. 1986

Price Rs. 50/-

Jacket designed by Shri. S. D. Phadnis Poona. Printed by Shri. Chandrakant Phadke at Kirloskar Press. Veer Savarkar Nagar. Poona-37 and published by Shri. D. D. Kulkarni. Shrividya Prakashan 250 C/16 Shaniwar Ashtabhuja Road. Poona 30.

PREFACE

Of all the competencies a teacher needs, the most neglected is that of evaluating student progress. It is not unusual to find teachers who lack a grasp of basic principles of evaluation or who lack the skill necessary to produce a good class-room test. Measurement of one kind or another takes place in our life frequently. A grocer weighs sugar to determine its price. A tailor measures certain dimensions of the body to produce a garment of proper-size. A motorist observes the speedometer to avoid exceeding the speed-limit. Gas and electricity lines are metered to determine the amount consumed. All these are instances of *quantitative* measures. There are also some less obvious instances of measurement. Ratings of popularity of cine-stars, ranking of girls in a beauty-contest are all examples where *qualitative* aspect is involved more than the quantitative aspect.

In a school situation, a teacher is concerned with both the *qualitative* and *quantitative* aspects of measurement. Unfortunately in our schools to-day, mostly the quantitative aspect is emphasized, with the result that all the programmes in a school are geared to written examinations. The process of education has become lop-sided and the final product of education is to-day creating anxiety among Indian leaders and educationists. This has been vividly illustrated on the cover page of this book.

The main theme which runs throughout the book is that evaluation is an integral part of the teaching-learning process and that it involves three fundamental steps : (i) Identifying and defining instructional objectives in behavioural terms, (ii) using suitable learning experiences and methods and (iii) constructing and selecting evaluation instruments which would most effectively appraise various learning outcomes, and using the results as feed-back material. In the absence of this integrated approach, we are sure to get frustrating results.

If you are enrolled in a first course in measurement and evaluation, if you want an orientation to this field, if you want to acquire basic understanding and skills in preparing various tools of evaluation, and if you are interested in the interpretation and application of results of measurement, the text is addressed to you. You may be a class-room teacher, or a supervisor or a student-teacher, the book will satisfy your needs. There are two simple rules which should guide your efforts. First, be sure, you understand each concept. Second, practice until the new concepts and processes become a secure part of your knowledge and skills. If you happen to be among those who have depended largely on rote memory to pass in an examination, now is the time to try something more effective. Blindly following a formula or memorizing a definition will not turn the trick. You must work with a concept until it is clear and until it becomes a part of your thinking. Remember that you learn what you think, feel and practice.

The general orientation of this book is functional. Throughout most of the chapters, there appear questions and exercises which challenge the reader to tease out implications, to search for illustrations from his own experience and to apply principles previously discussed. The functional aspect of this book can best be indicated by listing the learning outcomes which should result from its use. These outcomes are :

(i) An understanding of the interrelated nature of teaching, learning and evaluation.
(ii) Ability to define objectives in behavioural terms.
(iii) Skill for constructing various tools of evaluation for measuring specific learning outcomes.
(iv) Ability to use various tools.
(v) Ability to interpret results.
(vii) Ability to use the results as feed-back material.
(vii) An awareness of both the potentialities and limitations of various tools.
(viii) An awareness for improving the teaching-learning process.

It is not intended that these outcomes be attained by reading alone. Some of these refer to skills and attitude which can be developed through practical experience. The book, it is hoped, will help the reader change his attitude and acquire necessary skills.

I am most grateful to those writers whose writings have mostly shaped my approach to evaluation. A list of all these authors is given in the bibliography. My thanks are due to Shri D. D. Kulkarni of *Shri Vidya Prakashan* who has taken great pains in bringing out this book. I am also thankful to the Printers for taking keen interest in completing the work within the shortest span of time.

W. N. DANDEKAR

B/2 Parimal.
1239 Apte Road.
Pune 411 004
20-10-71

CONTENTS

Section I : Techniques of Evaluation

1 **The Concept of Evaluation** .. 9

Measurement and Evaluation, Principles of Evaluation, Evaluation Approach, Tools and Techniques of Evaluation, Implications of Evaluation Approach.

2 **Objectives and their Classification** .. 22

What is an Objective, Objectives and Learning Outcomes, Specifications of Objectives, Classification of Objectives, Taxonomy of Educational Objectives, Uses of Taxonomy.

3 **Instructional Objectives** .. 50

Objectives of English, Mathematics, Geography, Mother Tongue, General Science, Social Studies, Work Experience, Physical Education, Art and Music.

4 **Objectives of Teacher Education** .. 77

Knowledge, Application, Skills, Interest, Attitude, Appreciation.

5 **Learning Experiences** .. 84

Classification of Experiences, Criteria for choosing Experiences, Centres of Experiences, Activities pertaining to various School Subjects, Emotionalized Learning Experiences.

6 **The Unit Plan** .. 104

Meaning of the Term, Units in Specific Subjects, Steps in Unit Planning, Unit Plans in English, Science, Civics, Geography and Mathematics, Year's Plan.

7 **Unit Test** .. 119

Preparation of a Design, a Blue Print, Questions based on the Blue Print, Editing the Paper, Marking Scheme, Questionwise Analysis, Item-Analysis, A sample question paper in English, Scoring Key, Marking Scheme, Questionwise Analysis.

8 **Standardization of a Test** .. 135

Design of a Test, Preliminary Form of a Test, Item Analysis—difficulty value and discriminating power. Final form of a Test, Scoring Key, Establishing Norms, Validity and Reliability.

9 **Characteristics of Measuring Instruments** .. 151

Validity, Reliability, Objectivity, Adequacy, Usability, Interpretation of Test Scores, Raw Scores, Derived Scores, Types of Norms, Assigning Grades, Report Cards.

10 **Tools of Evaluation : I** (Quantitative Techniques) .. 179

Present day Examinations, Essay Type, Merits and Limitations, Improving the Essay Examination, Objective Type, Advantages and Limitations, Oral Examination, Performance Tests, Standardized Tests, Diagnostic Tests, Uses of Achievement Tests.

11 **Tools of Evaluation : II** (Quantitative Techniques) .. 203
Types of Objective Test Items, Suggestions for Writing Items, Completion Type, True-False Type, Matching Item, Master Matching Form, Multiple Choice Item, Interpretive Type.

12 **Tools of Evaluation : III** (Qualitative Techniques) .. 229
Observational Techniques : Rating Scales, Anecdotal Records, Check Lists, Cumulative Records; Self Report Techniques: Interview, Inventories; Projective Techniques, Sociometric Devices.

13 **Internal Assessment** .. 255
Importance, Areas and Plan of Internal Assessment, Weightage, Validity, Lesson Evaluation Form, Advantages and Limitations, Retrospect and Prospect.

Section II : Statistical Methods

14 **Frequency Distributions** .. 267
Need for Grouping, Preparation of a Frequency Table, Column Diagram, Frequency Polygon, Smoothing a Polygon.

15 **Measures of Central Tendency** .. 287
Mean and its Computation, Median, Special Cases of Median, Mode, Characteristics and Uses of these Measures.

16 **Measures of Variability** .. 311
Range, Quartile Deviation, Average Deviation, Standard Deviation, Interpretation, Uses and Limitations of these Measures.

17 **Measures of Relationship** .. 331
Meaning of Correlation, Degrees of Correlation, Rank Difference Correlation, Interpretation and Uses of Correlation Coefficient.

18 **The Normal Curve** .. 344
Meaning of Normal Distribution, Properties of the Normal Curve, Application of the Normal Curve, Skewness, Kurtosis, Other Types of Curves, Making Test Scores Comparable, Standard Scores, Percentiles and Percentile Ranks.

Appendix I and II .. 370
Objective based Test Items in Mathematics, English, Geography, Civics and Science.

Index .. 379

Bibliography .. 383

Section 1

Techniques of Evaluation

1. The Concept of Evaluation
2. Objectives and their Classification
3. Instructional Objectives
4. Objectives of Teacher Education
5. Learning Experiences
6. The Unit Plan
7. Unit Test
8. Standardization of a Test
9. Characteristics of Measuring Instruments
10. Tools of Evaluation : I (Quantitative Techniques)
11. Tools of Evaluation : II (Quantitative Techniques)
12. Tools of Evaluation : III (Qualitative Techniques)
13. Internal Assessment

1

THE CONCEPT OF EVALUATION

Meaning of Measurement and Evaluation

The word measurement means the act or the process of ascertaining the extent or quantity of something. Evaluation refers to the act or process of determining the value of something. Evaluation depends upon, but is not synonymous with measurement. Evaluation goes beyond measurement in answering the question: "Is the obtained measurement, desirable or undesirable?"

A familiar simple illustration may help clarify the distinction between measurement and evaluation. A teacher gives her class a spelling test of 40 words. She finds that the number of words spelled correctly ranges from 12 to 38, the average number correct being 26. Leela has a score of 17. This score denotes only measurement. Is this a good score or a poor score? If the teacher says, that 17 indeed is a poor score or says that in view of the limited verbal ability Leela is spelling about as well as can be expected, then the teacher is, *evaluating* the score or if the teacher says, Leela has the necessary ability, that she should have obtained a higher score but she does not work hard, then also it is a form of *evaluation*. When we assign grades such as excellent, good, fair, poor or high, average, low, we are evaluating the performance. Consider this analogy. If a scale shows that a man weighs 210 pounds, it is reporting a measurement. The fact may be unpleasant, but there is nothing personal about it and the scale is objectively reporting the facts. But when the doctor sees the report and says, "You are getting too fat Raju," he is making an evaluation. Evaluation is thus a qualitative judgment of how good or how satisfactory an individual's performance has been.

Sound evaluation is based upon the results of measurement. Considerable measurement is involved in purchasing a house.

The amount of floor space, the cost of brickwork, plumber and wiring are among the measurements to be considered. The judgments based on these measurements are an evaluation. Further, some intangible items would enter the picture; the style of the house, convenience of room arrangement, environment etc., would be taken into account although they are beyond the limits of precise measurement. Similarly some significant educational factors are beyond the limits of measurement e.g., tests do not measure drive or motivation for learning nor do they measure the student's attitude towards the school programme. Hence other types of data have to be used to evaluate the school programme.

From an educational point of view, *evaluation may be defined as a systematic process of determining the extent to which educational objectives are achieved by pupils.* There are two important aspects of this definition. (i) Evaluation implies a systematic process, which omits casual uncontrolled observation of pupils. (ii) Evaluation always assumes that educational objectives have been previously identified. Without previously determined objectives, it is impossible to judge the extent of progress. This definition indicates that evaluation is a much more comprehensive and inclusive term than measurement. Evaluation includes both *qualitative* and *quantitative* descriptions of pupil behaviour plus value judgments concerning the desirability of that behaviour. Measurement is limited to quantitative descriptions of pupil behaviour. It does not include qualitative descriptions nor does it imply judgments concerning the worth or value of the behaviour measured. Measurement may tell *how much;* evaluation may tell *how good.* We cannot say *how good* a student's progress is, unless we know *how much* he has achieved. The main emphasis in evaluation is on the extent to which educational goals are achieved. We must be able to answer a number of questions through the process of evaluation. In the teaching of English a teacher has constantly to answer such questions : Is Rajesh able to read fluently ? Is his handwriting legible ? Can Madhuri express her ideas in speech or in writing ? Is she interested in the study of English ? Has she developed good study habits ? Can she use a dictionary ? Can she participate in discussion ? These are typical questions which teachers have to answer. A variety of methods is therefore necessary and a sound evaluation programme will include both measurement and non-measurement techniques.

General Principles of Evaluation

Evaluation is sometimes viewed as merely a collection of techniques for obtaining information about pupils. However, evaluation is not a mere *collection* of techniques. Evaluation is a *process*. This process like the process of teaching is most effective when it is based on sound principles. The following principles provide direction to the process and serve as criteria for appraising the effectiveness of specific procedures.

1. No evaluation device should be selected or developed until the purposes of evaluation have been carefully defined. In assessing a pupil's progress, we must first of all identify and define educational objectives. By identifying educational objectives, we know what is to be evaluated. The achievements of India's Five-Year Plans are evaluated from time to time. While evaluating the plan, the research workers or the critics take into consideration the objectives of the Five-Year Plan, such as increasing the agricultural production, raising the per-capita income, establishing basic industries, utilizing available resources, etc. The critic tries to see how far these objectives have been achieved. Similarly in evaluating the success of any scheme in any field, one has to identify and to define the objectives of the scheme. In the school situation, the teacher must clearly know the objectives of teaching English, or Mathematics, or Science, before she proceeds to evaluate the student's progress in any of these subjects.

2. Evaluation techniques should be selected in terms of the purpose to be served. There are various techniques of evaluation; one technique may be appropriate for some purpose but inappropriate for others, for example, a written examination will not be an appropriate tool for assessing a student's skill in handling the apparatus or his ability for social adjustment. You will have to devise some other suitable tool for evaluating these latter aspects.

3. Comprehensive evaluation requires a variety of evaluation techniques. No single evaluation technique is adequate for

appraising pupil's progress toward all of the important outcomes of instruction. To get a more complete picture of pupil's achievement, we need to combine the results from a variety of techniques.

4. Proper use of evaluation techniques requires an awareness of their limitations as well as of their strengths. Some of the evaluation techniques (such as standardized tests of intelligence and achievement) are fairly good measuring instruments, while a few tools (such as rating scales) are crude instruments. All of these tools are subject to one or more types of error. For example, the scores on objective tests are influenced by guessing while those on essay tests are influenced by the subjective judgment of the examiner. Another type of error arises from improper interpretation of evaluation results. People are unwilling to recognize the limitations in the evaluation instruments and they attribute to these instruments a precision which they simply do not possess. Jayant has a score of 82 and Madhuri has a score of 83 on an achievement test. It will be risky and also unfair to make a distinction between the achievement of these two students on the basis of a difference of 1 score-point. As a guiding principle, it will be helpful to keep in mind that the cruder the instrument, the greater the limitations and consequently the more skill needed in its use.

5. Evaluation is a means to an end and not an end in itself. The use of evaluation techniques implies that some useful purpose will be served and that the teacher is clearly aware of the purpose. To blindly gather data about pupils and then to file the information away in the hope that it will some day prove useful is a waste of time and effort. Some schools collect information about children's weight, height, eye-sight etc., but no follow-up work is done. It means that this collection is done just as a ritual to satisfy some inspecting officer. Many activities are done in this way without due regard to the purpose with the result that there is a huge wastage of time and effort.

The Evaluation Approach

Evaluation is a comprehensive term. It has a bearing on all the stages of the process of education, beginning with the objectives, planning of the syllabus, teaching-learning process and testing.

The main purpose of class-room teaching is to change pupil-behaviour in desired directions. When viewed in this light evaluation becomes an integral part of the teaching-learning process. The 'desired directions' are the educational objectives set forth by the framers of curriculum. Evaluation is the process of determining the extent to which these objectives are being achieved. For understanding the relation between evaluation and the teaching-learning process, let us study the various steps involved in the process of *Education* (Fig. 1.1).

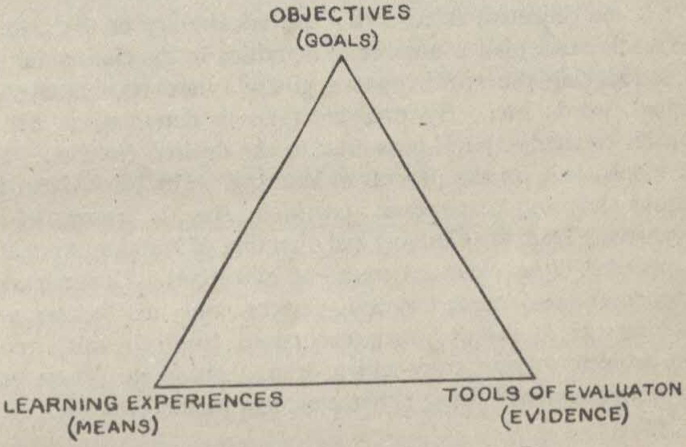


Fig. 1.1 : The Triangle of Evaluation.

(i) *Identifying and defining Objectives* : The first step in both teaching and evaluation is that of determining the learning outcomes to be expected from class-room instruction. What abilities and skills should be developed when a child studies the mother tongue and arithmetic for one year? What type of interests and attitudes should be developed among the children? What changes in habits of thinking and doing should take place? In short what behaviour changes do we expect in the children?

Only by identifying objectives and stating them clearly in terms of specific behaviour, we can provide direction to the teaching process.

(ii) *Planning Learning Experiences* : We can give direction to the process of teaching if we identify and define the objectives. For example, one of the objectives of teaching English is:

"The student reads English with comprehension."

The teacher will have to provide opportunities for silent reading in the class-room. She can then test whether the students read with or without comprehension. The teacher will gradually increase the length of the passages for reading and maintain the same time-limit, thus encouraging the students to improve their speed of reading.

If the objective is to enrich the vocabulary of the students, the teacher can plan a number of activities in the class-room such as introducing the spelling-game, giving synonyms, antonyms of various words etc. By organizing these activities, it will be possible to change pupil-behaviour in the desired direction. Here the emphasis is on the *process* of learning. The teacher can plan various learning experience provided she is aware of the objectives. Thus the planning and directing of learning experiences is preceded by a clear statement of objectives. Under learning experiences are included various devices, aids and techniques of teaching such as use of a map, excursions, library reading, recitation, problem-solving, story-telling, dramatization etc. These learning experiences will bring about a change in the behaviour of the student.

(iii) *Providing evidence through various tools of evaluation*: The product of learning will be seen in the knowledge and skill acquired or in the interest and attitude developed by the pupil. The teacher has then to use some tools for measuring the change in behaviour. In the above illustration the teacher can see whether the speed of reading and vocabulary have been improved, and whether the student has developed interest in reading. Evaluation thus helps in determining the extent to which the objectives are being attained.

(iv) *Using the results of evaluation as feed back for improving the teaching-learning process*: If the teacher finds that there is neither any improvement in the reading ability of the student nor any interest has been developed in reading, she can make use of this information for improving her methods.

and for improving the learning experiences. Teacher can also reconsider the objectives if they are found to be impracticable.

We thus find an intimate relation between the objectives of teaching, learning experiences and the tools of evaluation. This relationship can be symbolically represented as shown in Fig. 1.1.

Objectives remain central to both learning experiences and evaluation. Learning experiences should be objective-based, and evaluation should also be objective-based. Learning experiences are planned in terms of objectives. Evaluation tools are to be selected with reference to the objectives. A written examination for example, should not be used for assessing the interest developed by a student in reading or for assessing the laboratory skills of a student in science. Table 1.1 shows how learning experiences and evaluation are both objective-based.

Tools and Techniques of Evaluation

One of the distinctive features of the evaluation process is the use of a wide variety of procedures. As indicated earlier, some of these may be classified as *quantitative* techniques because the results can be reduced to numerical scores. Other means of appraisal are classified as *qualitative* techniques because their results can be expressed only in verbal descriptions. Information for evaluation may be obtained by presenting an individual with a given set of tasks to perform or by asking him questions about himself or by asking other persons to observe and to judge his behaviour. These general methods of obtaining data help us in classifying the various techniques of evaluation. They are commonly referred to as (i) *testing procedures,* (ii) *self-report techniques,* (iii) *observational techniques* and (iv) *projective techniques.*

1. *Testing Procedures* : A test is merely a series of tasks which are used to measure a sample of a person's behaviour at a given time. The most common tests used in schools are achievement tests. They may be oral, practical or written. Written tests may be further classified in a number of ways. They may be essay type or objective type; they may be standardized or informal; they may be survey tests or diagnostic tests. We shall consider these testing procedures in chapters 8, 10 and 11.

Table 1.1 : Objective-based instruction and evaluation

Objective	Learning Experiences	Evaluation Tools
(1) Skill in handling apparatus in a laboratory	(i) Asking students to perform experiments (ii) Arranging workshop practice	Practical examination
(2) Developing oral expression	(i) Carrying discussions in the class-room (ii) Asking students to participate in elocution competition	(i) Observation of students in various suitable situations (ii) Oral examination
(3) Skill in reading maps	(i) Using maps in the class-room (ii) Asking students to use maps	(i) Oral examination for map-reading (ii) Asking students to draw outline maps and filling the same
(4) Ability to write correct English	(i) Using various structures, words, phrases etc. (ii) Asking students to use the structures	Written examination
(5) Developing understanding in Mathematics	Asking students to classify the data, to detect errors, to find relationship among the data	(i) Oral examination (ii) Written examination
(6) Developing interest in science	(i) Allowing students to work in a laboratory and in a workshop (ii) Organizing science club-activities (iii) Reading books on science	Observation of students in various situations such as a workshop, a science-club etc.

2. *Self-Report Techniques* : Every individual has a wealth of information about himself. Some of this information such as how he feels about certain situations which activities interest him most or what personal problems are of greatest concern to him, can only be obtained directly from the individual. For evaluation purposes, this information is usually obtained by an *interview* or a *questionnaire*. The interview obviously involves a face to face relationship between the interviewer and another individual. When used as a self-report technique, the interviewer is mainly interested in securing information concerning the individual's attitudes, opinions, interests and the like.

The questionnaire method of obtaining information from individuals is most commonly used in systematic attempts to evaluate interests, attitudes, and other aspects of personal and social adjustments. It has an advantage over the interview method in that it can be used in group situations. Each individual is presented with a series of questions or statements to which he must respond by answering 'yes' or 'no', 'agree or disagree' or in some other manner indicate his feelings and opinions. Under this category we get *personality inventories*, and *interest inventories*. Since the usefulness of both the interview and the questionnaire depends upon the willingness of the individual to give honest answers, they are decidedly limited for purposes of evaluating pupil progress. They are probably most useful for guidance purposes, where the individual's desire to understood himself and to make wise future plans encourages him to describe himself as he actually is rather than as he would like to be. We shall describe these techniques in chapter 12.

3. *Observational Techniques* : Reliable information about an individual's typical, or usual behaviour is best obtained from persons who have observed him in a variety of situations. In the case of pupils, both teachers and the pupils' peers have numerous opportunities to observe their typical behaviour. The various observational techniques are merely systematic methods of recording these observations for purposes of evaluation. *Anecdotal records, checklists, rating scales* and *sociometric techniques* are included in this category. These will be discussed in chapter 12.

4. *Projective Techniques*: These techniques provide another method of evaluating personal-social adjustment with which the class-room teacher should be familiar. The self-report techniques are structured, that is, they tend to fix the direction of response of the pupil. In contrast to these structured inventories, projective techniques provide almost complete freedom of response. Some of the techniques used under this category are *sentence-completion, doll-play, perception of ink-blots, interpretation of pictures* etc.

We can classify the various tools of evaluation as follows :

Table 1.2 : Tools and Techniques of Evaluation

1 **Testing Procedures**	2 **Self-report Techniques**	3 **Observational Techniques**	4 **Projective Techniques**
(i) Written examination	(i) Interview	(i) Check-list	(i) Sentence completion
(ii) Oral examination	(ii) Interest inventory	(ii) Anecdotal record	(ii) Doll-play
(iii) Practical examination	(iii) Personality inventory	(iii) Rating scale	(iii) Story completion
(iv) Pupil's products	(iv) Questionnaire	(iv) Sociometric technique	(iv) Ink-blot
(v) Standardized tests	(v) Problem check-list	(v) Cumulative record	
(vi) Diagnostic Tests			

Implications of the Evaluation Approach

Evaluation being a dynamic process has a profound influence upon the different aspects of education. Its influence can easily be felt in the areas of

(i) learning (ii) teaching (iii) syllabus (iv) text-books and (v) testing procedures.

(a) *Influence on Learning*: (i) It ensures regular study habits because it is a continuous process. (ii) It deemphasizes memory as the new approach puts emphasis on the achievement of all the objectives of instruction and does not take into consideration only the knowledge or information aspect. Pupils strive hard for the achievement of abilities such as understanding, appreciation, skills etc. (iii) It discourages selective study. Wide coverage of content in the question papers necessitates that the pupils prepare the entire range of the syllabus.

(b) *Influence on Teaching* : (i) At present the lecture method occupies the most prominent place in the whole scheme of teaching. Teachers are the stockists who supply information to the customers. This attitude will change and the teacher will use such methods as will help develop the various abilities of children. (ii) It will involve greater planning and preparation on the part of the teacher for the achievement of pre-determined objectives of instruction. (iii) It helps the teacher to accelerate the progress of the child because of diagnostic and remedial teaching.

(c) *Influence on Syllabus* : (i) This approach necessitates recasting of the syllabus. Instructional objectives are given at the beginning of the course-content. (ii) Only such topics are retained in the syllabus as will help fulfil the instructional objectives. (iii) Scope of the various topics is clearly mentioned.

(d) *Influence on the Text-Books* : (i) Change in methods of teaching and reorganization of the syllabus would naturally influence the format of text-books. At present text-books are not generally written with the specific purpose of achieving instructional objectives. They are merely a compendium of information. (ii) Because of the evaluation approach text-books will be so modified as to take up the objective-based approach. They will incorporate a variety of activities for achieving the objectives.

(e) *Influence on Testing* : (i) At present the emphasis is laid on the written examination. This emphasis will be shifted and various techniques will be used for assessing the progress of the child. (ii) There will be a reform in the examination system. Objective-based questions will be set. (iii) Assessment will be objectively done. (iv) Results will be interpreted in a more valid and reliable way.

Evaluation approach thus aims at the establishment of close inter-dependence between educational objectives, learning experiences and appraisal of the total school programme so that each may modify itself and grow in correspondence with the rest.

What are the Uses of Evaluation ?

(1) Evaluation is useful in revealing the stage at which pupils have arrived in the learning process. It permits the teacher to appraise the extent to which pupils are progressing towards the goals of instruction. She can plan her programme with reference to the progress of her students. (2). Evaluation properly planned and conducted motivates pupils towards better attainment and growth. It serves as a stimulus for greater effort on the part of the students. (3) Evaluation is useful to test the efficiency of the teachers in teaching. It helps us to test the effectiveness of instruction and of the class-room activities. (4) It serves the basis of reporting to parents about the progress of their wards in the school. It provides the teacher with an objective and comprehensive picture of each pupil's progress. (5) The results of evaluation procedures are especially useful for guidance and counselling. (6) A comprehensive evaluation programme helps the Headmaster also. He is able to judge the extent to which the objectives of the school are being achieved, to identify strengths and weaknesses in the curriculum and to appraise the various activities in the school.

An Innocent Question

Miss Kaveri, who teaches geography and science to standards VII and VIII says, "We use periodical tests for assessing the progress of girls in our school; we have also improved the nature of these written tests. What then is the difference between evaluation and these periodical tests ? "

Remember that written tests are just one of the many tools of evaluation. Testing is not synonymous with evaluation. Written tests give you a quantitative description of pupil's progress. The written tests do not give any answers to such questions as :

Have the girls acquired any laboratory skills ?
Have they developed the scientific attitude ?

Are they interested in the study of birds and other creatures ?
Do they go out rambling for observing the habits of birds ?
Can they read maps, charts and graphs showing weather changes ?

Answers to all these questions are not obtained from the written examination. A comprehensive scheme of evaluation is therefore necessary. Use of various tools will enable the teacher to answer the various questions mentioned above.

Evaluation = Quantitative description of pupils + Qualitative description of pupils + Value judgments about the above two descriptions.

FOR FURTHER STUDY

1. Bring out clearly the difference between measurement and evaluation.
2. Discuss the general principles of evaluation.
3. Bring out clearly the relation between objectives, learning experiences and tools of evaluation.
4. Explain the terms : Objective-based learning experiences and objective-based evaluation tools. Illustrate with suitable examples from two school subjects.
5. Discuss the various tools and techniques of evaluation. How will you classify them ?
6. How does the evaluation approach influence the different aspects of education ?
7. State and explain the various steps involved in the process of evaluation.

.8 How does the evaluation approach help the Headmaster in the appraisal of the total school programmes ?

*

371.26
DAN

S.C.E.R.T., West Bengal
Date 24 10 89
Acc. No. 4727

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

2

OBJECTIVES AND THEIR CLASSIFICATION

The whole process of education is geared to the objectives. You cannot frame the syllabus if the objectives are not made known to you. Teachers will fail to adopt suitable teaching methods in the absence of the knowledge of objectives. Evaluation will be meaningless if it is done without reference to instructional objectives.

When asked, " What is your objective of teaching the mother-tongue ? " Meena, a smart teacher, answers promptly, " I intend to complete the text-book prescribed by Government, wherein there are 24 prose lessons and 8 poems. " Because of this attitude all her efforts are directed towards completing the text-book. The pupils learn by heart a few answers, reproduce them in the answer-books and get through the examination.

Dinesh who has passed the S. S. C. examination has applied for a clerical job in a bank. He has been called for an interview. The interviewing officer has written the following remarks after interviewing Dinesh :

(i) Handwriting of Dinesh is not legible.
(ii) He has failed to express his ideas logically.
(iii) He could not tell the names of any books he has read.
(iv) He has not developed any interest in reading.
(v) It appears that he does not read the newspapers daily.
(vi) He committed a number of mistakes when he was given some dictation.
(vii) He could not discriminate between facts and opinions.
(viii) He could not convince others with his view-point.
(ix) His vocabulary was very poor.
(x) His pronunciation was defective.

Remarks : The candidate need not therefore be selected.

How is it that Dinesh could get through the examination in spite of the above mentioned deficiencies ? It means the teacher did not attend to such things as pronunciation, handwriting, dictation, oral and written expression, reading of books etc. A student feels that without achieving these skills he can get through the examination. Dinesh is not the only one of this type. He is the representative of a large number of students of this type. This happens simply because teachers neglect many of the objectives of teaching the subject and simply try to teach the text-book. They forget that the text-book is a means and not an end in itself. A teacher may be justified if she says that she can achieve the objectives without the use of a text-book. Such teachers however, are rare.

What is an Objective?

An objective is a point or an end-view of something towards which action is directed. It gives direction to the programme that one wants to undertake.

While preparing the draft of India's Five-Year Plan, the planning commission sets before it some objectives that should be achieved during the plan-period. A few of the objectives can be stated as follows:

(a) To increase agricultural production.
(b) To raise per-capita income.
(c) To establish basic industries.
(d) To utilize available resources.
(e) To establish a net-work of communications.
(f) To provide schooling facilities to all children of the age group 6-11.
(g) To provide medical facilities.
(h) To banish poverty.

While evaluating the plan, one has to take into consideration the objectives as the frame of reference. If we disregard the objectives the evaluation done will mislead people. A student while reviewing the plan said, "The number of cars in the city of Bombay has increased from 37,000 to 54,000 during the III Five-Year Plan. Production of terene was doubled while the sale of transistors was trebled in Bombay. Hence we have succeeded in the implementation of the plan." Another student said, "The per-capita income in U.S.A. is Rs. 14,000 per year, that in Japan is Rs. 4,000 while ours is only Rs. 460. Hence we have failed in the implementation of the plan."

If we carefully study the above arguments, we shall find that the approach of the critics is not correct. The critics have not taken into consideration the objectives set forth by the planning commission. Our objective was not to raise the per-capita income to the level of U. S. A. or Japan. Increase in the number of motor cars cannot be considered as the *true* index of achievement. Reference point should always be the objectives.

The same principle applies to the school-situation. The secondary school programme extends over a period of seven years. It is as if *a seven-year plan* similar to India's Five-Year Plan. Just as the planning commission predicts the picture of India's development through a statement of objectives, similarly the teachers must have a clear picture of the child's development during the seven-year-plan-period. This is done by formulating statements of objectives.

A guardian approaches the Headmaster and says, "I am admitting my son Rajesh in standard V of your school. He will be there under your control for a period of seven years. How will he look like at the end of this period?"

The teacher says, "Our aim is to bring about an all-round development of the child."

The guardian is puzzled by this jargon. He does not understand the implications of the term "all round development."

The teacher says, "The teacher is an architect and he gives proper shape to the developing child who is like a lump of clay."

The guardian is still further puzzled by this jugglery of words. He says, "Point out to me the change in behaviour that you intend to bring about in the child."

The guardian's plain questioning triggered the mental process of the Headmaster. He discussed the matter with his colleagues, who had also used the same jugglery of words. All of them set out to analyse the term "All round development." They came out with the following description.

Fig. 2.1 : No learning experiences

Behaviour after providing learning experiences

(1) He can read and write well.
(2) He can play games.
(3) He can express his views.
(4) He can solve problems in daily life situations.
(5) He works hard.
(6) He is not disappointed by failures.
(7) He shows sympathy for the poor.
(8) He takes interest in games.
(9) He has developed scientific attitude.
(10) He has the qualities of a good citizen.

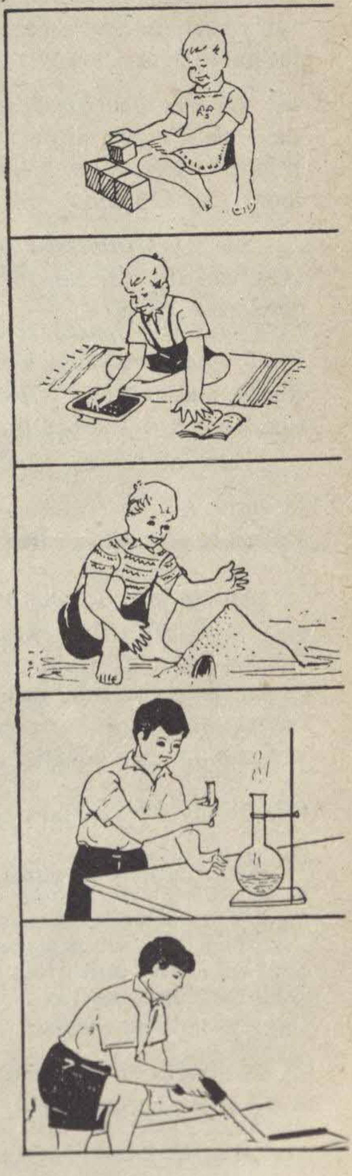


Fig. 2.2 : Changes in child-behavi

The objective ' all round development ' was not meaningful. In order to make it meaningful, we have tried to translate the term ' all round development ' in fig. 2.2. Some of the statements in the above figure however need further classification, for example :

He has developed scientific attitude : What is meant by the term scientific attitude ? How does a child with this attitude behave ? How does he differ from the child who has not developed this attitude ?

He takes interest : How is interest seen in his behaviour ? How can the teachers or parents recognize that the child has developed interest ?

He can read and write well : What are the marks of good reading and writing ? How does such a child differ from the one who has not developed this ability ?

These objectives therefore need further specification.

Statement of an Objective

An objective should be stated in specific terms. While spelling out the objective we are concerned with the kind of change that we intend to bring about. Since an objective is an end-view of something, we shall mainly emphasize the kind of change to be brought about. If a patient is under the treatment of a doctor, how can you say whether the doctor has achieved the objective ?

Objective: *To improve the health of the patient.*

Specifications in terms of behaviour changes :

(i) There is a gradual increase in weight.
(ii) The patient enjoys sound sleep.
(iii) The patient has a good appetite.
(iv) He feels fresh.
(v) He has normal temperature.

Achievement of the objective will be judged with reference to the above specifications.

' Banish Poverty ' is a popular slogan in India to-day. If this is the objective, it can be spelled out as follows :

Objective: *To banish poverty.*

Specifications:

(1) Per-capita income will be increased from Rs. 460 to Rs. 700 during the plan-period.
(2) Fifty per cent of the slums will be removed.
(3) Minimum wages of farm-labourers will be increased by 40 per cent.
(4) Employment opportunities will be created for the semi-skilled and unskilled workers.
(5) Ceiling on land-holdings will be imposed.
(6) Landless labourers will be provided with land.
(7) Ceiling on accumulation of wealth will be imposed.
(8) The gap between the *haves* and the *have-nots* will be reduced.

This list of specifications can further be increased. The achievement of the objective is to be judged with reference to these specifications.

Meaning of the term Educational Objective

Educational objectives are goals towards which pupils progress. They are the end results of learning stated in terms of changes in pupil behaviour. It may be defined as a desired change in behaviour. The term *behaviour* as used here, refers to mental, emotional and physical reactions. Thus an increase in knowledge, a broadening of understanding, an improvement in physical skill, a shifting of attitude and a deepening of appreciation are all classified as changes in behaviour (Fig. 2.2). If education imparted is effective, then the student will behave differently from the way he did before he came to school. The pupil knows something of which he was ignorant before. He understands somethings which he did not understand before. He can solve problems which he could not do before. He revises his attitudes desirably towards things. If Sheila has learned science for a number of years in a school, we must be able to show that she behaves differently from the way she did before she learned science.

It will be seen from Table No. 2.1 that Sheila behaves differently after learning science. How did this change come about? Obviously the teacher must have provided adequate opportunities to Sheila in the class-room. These opportunities are termed as *learning experiences*. A teacher has to arrange these experiences in such a way that they will lead the pupils towards attainment of objectives.

Table 2.1 : Sheila's Behaviour

Before learning science	After learning science
(i) She was ignorant about scientific knowledge	(i) She has now a knowledge of scientific principles
(ii) She did not think critically	(ii) She now thinks critically
(iii) She had no interest in scientific literature	(iii) She is now interested in scientific literature
(iv) She had no skill of handling any instruments	(iv) She has now acquired skill in handling instruments
(v) She used to pass judgments without waiting for adequate evidence	(v) She does not now jump to conclusions without getting evidence
(vi) She had a dislike for scientists	(vi) She now appreciates the work of scientists

Objectives and the Learning Outcomes

The objectives of teaching any subject should be laid down in unambiguous and specific terms. Overall goals like 'developing the child into an efficient citizen' are too vague and general. The teacher neither knows the specific change in the behaviour of the child nor the process of bringing about the desired change. Every statement of an objective should describe the kind of change or growth expected in the child when the objective is realized. When an objective is realized, we call it the *learning outcome* or *the product* which is expressed in terms of change in behaviour. It then becomes meaningful to the teacher.

A clearly defined objective has two dimensions. One dimension describes the kind of change or modification to be brought about in the student and the second part describes the content through which the change is to be brought about. The first part may be called the *modification* part or the *product* and the second one as the *content* part.

(a) *Example* : The pupil reads with comprehension an unseen passage.

Here the words 'reads with comprehension' denote the modification part. It is the learning outcome, while the words 'an unseen passage' denote the content part.

(b) *Example* : 'The pupil writes English correctly.

Here the words 'writes correctly' denote the modification part while the words 'writes English' denote the content part.

It is important to keep in mind that while formulating an objective we are more concerned with the product or the learning outcome or with the modification part rather than with the process of learning. The process may vary from school to school as it depends upon the facilities available in the school. An educational objective is not something that the teacher does, it is not the same as course content. Consider the following statements :

(i) To illustrate through films the life of Tibetians.
(ii) To demonstrate the various types of thermometers.
(iii) To demonstrate the skills in swimming.
(iv) To show how maturation influences readiness in learning.
(v) To teach lesson XIV from the language text.

The above statements cannot be called as ***educational objectives***. They indicate the process which the teacher wants to follow. These are statements about the means and not about the ends. Hence while stating an educational objective we must keep in mind that we are interested in making a list of expected learning outcomes. We are not identifying the subject matter but the reaction the pupils will make to it. We are not listing learning experiences of the pupils but are listing the changes in pupils' behaviour; we are not describing what we intend to do in the class-room but we are making a list of the expected results of instruction. In stating the objective, the main concern is the pupil and what he will be like at the end of the teaching-learning process.

Stating objectives in terms of product rather than the process admitted is easier said than done. As stated in Meena's case at the beginning of this chapter, most of us are more concerned with the content of the course and the ongoing process in the class-room rather than with the expected changes in pupil-behaviour. We focus our attention on the content and the class-room instruction. We have to shift this focus from the content to the pupil. We must continually ask ourselves : " What should the pupils be able to do at the end of the course that they could not do at the beginning ? " We are then in a position to define the educational objectives in terms of product.

Specifications of Objectives

In order to clarify and understand an objective it is further defined into its elemental parts. These elements are called *specifications*. They delimit the scope of each objective. Specifications are given in terms of learning outcomes.

Objective : *To develop the ability to read with comprehension.*

In order to spell out the learning outcomes the teacher asks a question to himself :

(i) What is it that a child having this ability does ?
(ii) How does this child differ from the one who has not developed this ability ?

Specifications or *Learning Outcomes* :

(a) The pupil grasps the meanings of words, phrases and sentences.
(b) The pupil locates key words, phrases and sentences in a passage.
(c) He locates significant facts and ideas.
(d) He discriminates between statements of facts and of weiws.
(e) He identifies relationships between ideas, events, facts etc.
(f) He reads with reasonable speed and flow.

By translating the objective in terms of learning outcomes we find that the various aspects of this ability namely ' Reading Comprehension ' can be evaluated subsequently.

They indicate what a pupil can do when he achieves the general objective. Let us consider another objective in the teaching of the mother tongue.

Objective : *To develop the power of written expression.*

What specific behaviour should be accepted as evidence that a pupil has developed this ability ? How does he differ from the one who has not developed this ability ?

Learning Outcomes :

(a) The pupil uses appropriate words, phrases and sentences.
(b) He uses varied sentences.
(c) He spells the words correctly.
(d) He uses capital letters and marks of punctuation correctly.
(e) He organizes his ideas logically.
(f) He presents his ideas logically.

This list can further be elaborated. Through these learning outcomes we can say that the objective is being achieved.

When objectives refer only to the mental process, they may not convey the same meaning to all teachers. Consider the following statements of objectives :

To develop the reasoning ability of students.
To foster interest in mathematics.
To develop scientific attitude.
To develop in children a sense of civic responsibility.
To develop the power of understanding.
To develop the power of oral expression.

In the above statements various mental processes such as thinking, understanding, attitude formation, interest, power of expression etc., are mentioned. Different readers will not catch the same meaning of these terms. Therefore the statement of an objective should be further specified so that the mental process of thinking or reasoning will be reflected in the behaviour of the child. For example how does the child with a scientific attitude behave ? How is this attitude reflected in the behaviour ? Similarly how does a child who has developed the power of oral expression differ from the one who has not developed this ability ?

The objective should therefore be specified in terms of observable change in pupil behaviour. For example :

Objective : *To develop scientific attitude.*

Specifications in terms of behaviour changes : This attitude is reflected in such qualities as open-mindedness, intellectual honesty and suspended judgment. A pupil who has developed this attitude :

(i) suspends judgment until all of the facts are available.
(ii) does not believe in superstitions.
(iii) seeks cause and effect relationship in the data.
(iv) make interpretations which are free from bias.
(v) shows willingness to consider new interpretations and thus to revise the old ones.
(vi) does not accept opinions without proof of justification.
(vii) indicates confidence in the data obtained by scientific procedures.

We can evaluate the pupil's progress by observing his behaviour in various situations mentioned above and pass a judgment regarding the development of scientific attitude.

Objective: *To develop in children a sense of civic responsibility.*

Specifications in terms of behaviour changes : This quality is reflected in one's day-to-day behaviour in such matters as :

(i) He respects the law.
(ii) He accepts responsibility in social life.
(iii) He is sensitive to disparities in society.
(iv) He accepts his duties.
(v) He does not waste national wealth.
(vi) He listens to other's view point.
(vii) He gives preference to interest of the society when it conflicts with his own interest.
(viii) He stands in a queue and waits till his turn comes.
(ix) He does not use unfair means (e.g. giving or accepting bribe) in doing his duties.
(x) He believes in the rights of others.

We can observe the behaviour of the students in various situations and then say whether the objective is being fulfilled.

In a teachers college, one of the objectives of the training programme is :

Objective: *To develop ability in teaching.*

Specificatons in terms of behaviour changes : How does a teacher who has developed this ability differ from the one who hasn't this ability ? What is it that such a teacher does in the classroom ?

(i) The teacher knows the subject matter well.
(ii) He introduces the topic methodically.
(iii) He creates readiness among the pupils.
(iv) His narration is effective.
(v) He makes the students participate in the discussion.
(vi) He is confident.
(vii) He pays individual attention.
(viii) He asks thought-provoking questions while teaching.
(ix) He lays emphasis on activity method.
(x) He is not baffled by questions asked by students.
(xi) Pupils appear to be engrossed in the learning process.
(xii) He has good control over the class.
(xiii) The teaching is objective-oriented.
(xiv) He uses appropriate learning situations.
(xv) He does not feel bored but enjoys teaching.
(xvi) He is resourceful.
(xvii) He inspires pupils for further study.

We can elaborate this list further. While judging the teacher's ability to teach, our frame of reference will be the above points which can be observed in the class-room. When method-masters in a training college observe lessons, they should have some common frame of reference for evaluating the lesson.

It will be seen from various examples given above, that an objective which has no clarity in meaning because of the jargon, becomes immediately concrete, vivid and meaningful even to a layman, when it is stated in terms of behaviour-changes. We thus get an *operational definition* of an objective through specifications.

Criteria for formulating Objectives and their Specifications

We have so far discussed a few objectives with their specifications in terms of change in behaviour. On the basis of this discussion, we can now state the various criteria for writing down specifications:

(1) An objective should be stated in terms of pupil-behaviour. A good statement of an objective has two parts: The modification part and the content part, for example:

'The pupil reads with comprehension an unseen passage'.

Here the words 'reads with comprehension' denote the modification part, while the words 'an unseen passage' denote the content part.

(2) It should be in terms of observable change in pupil-behaviour. Objectives stated in terms of mental processes such as developing reasoning, thinking, understanding, appreciation, interest are not meaningful to the reader. They do not convey the same meaning to all readers. Such statement as 'the pupil understands, he appreciates, he takes interest, he thinks etc.,' must be further translated if they are to provide a basis for evaluation. The evaluator must know exactly what behaviour a pupil should be able to exhibit if he 'understands' or if he 'knows' or if he 'appreciates,' for example:

(i) The child tells the meaning of the terms 'noun and pronoun.'
(ii) He recognizes nouns and pronouns in a sentence.
(iii) He uses nouns and pronouns in sentences of his own.
(iv) He locates errors in the use of nouns and pronouns.

The objective thus becomes measurable. The statement should include an action verb such as *describe, explain, use, compare* rather than a verb referring to a mental state such as *understand, know, appreciate* etc.

(3) The terminology of the statement should be simple enough to be appreciated by the pupils, teachers and parents, for example:

'The pupil uses appropriate words' in sentences.'

Here the statement is quite simple.

(4) The statement should refer to only one objective. More than one objective should not be expressed in the statement, since in such cases it becomes difficult to identify the learning outcomes. For example, 'the pupil should be able to translate English into Hindi and Hindi into English with correct use of idioms' is not as clear a statement as when these are stated as three separate objectives.

(5) The statement should indicate a worthwhile objective. It should be achievable. For example in the teaching of music in secondary schools the objective should not be 'preparing professional musicians.' It is not a worthwhile objective since you cannot expect all pupils to be professional musicians. But you can state the objectives this way:

To acquaint the pupils with musical instruments.
To enable the pupils to handle musical instruments.
To enable the pupils to recite poems with rhythm.

(6) The statement should refer to the expected level of attainment, e.g. in skill subjects such as typing, the level of attainment should be specified:

The pupil types accurately 40 words per minute.
The pupil recites 8 lines of a poem.
The pupil solves 20 addition problems in 10 minutes.

If we are unable to define an objective in terms of specific learning outcomes as mentioned above, the objective as stated is meaningless and should be eliminated or revised.

Classification of Objectives

Objectives can be classified in a number of ways such as general versus specific objectives, tangible versus intangible objectives, ultimate versus immediate objectives. Following illustrations will make the difference clear.

1 (a) *General Objectives*: Ability to understand, ability to solve problems, ability to express ideas.

(b) *Specific Objectives*: Ability to understand meaning of words, ability to add numbers, ability to use appropriate words in sentences.

2 (a) *Tangible*: Ability to read a thermometer, ability to recite a poem, ability to type 40 words a minute.

(b) *Less tangible*: Developing interest in mathematics, developing appreciation for beauty, ability to adjust in society.

3 (a) *Ultimate*: Developing an all-round personality, developing good citizenship, developing critical thinking etc.

(b) *Immediate* : Ability to recognize the meaning of words, to use words in sentences, to locate key words and sentences in a passage etc.

Ultimate or general objectives are essentially the same at all levels from Kindergarten to grade XI. Specific or immediate objectives are the narrower day-to-day goals which ultimately lead to the achievement of general objectives (Fig. 2.3).

Achievement of general objectives depends upon the contribution of each specific objective. If the specific objectives run in various directions there will be chaos and the general or ultimate objectives will not be achieved.

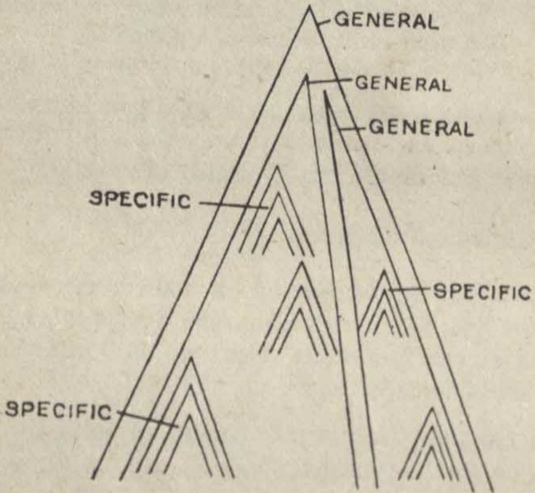


Fig. 2.3 : General and Specific Objectives.

Consider the following statements of objectives :

1. To bring about an all-round development of the child
2. To develop the reading ability.
3. To develop skill in reading.
4. To develop scientific attitude.
5. To develop a sense of civic responsibility.

It will be seen that the first statement denotes the type of ultimate or general objective while the remaining ones are specific objectives. They can be achieved in a limited span of time. But all these objectives point towards the ultimate objective. Each of the four specific objectives above can further be made more specific, for example, consider how objective No. 3 above can further be specified:

(a) To enable the student to read with reasonable speed and flow.
(b) To enable him to grasp the meanings of words.
(c) To enable him to locate key words, phrases and sentences.
(d) To enable him to locate important ideas and facts.
(e) To enable him to identify relationship between ideas, events, facts etc.

These specified or immediate objectives are pointing in the same direction. They do not conflict with one another since their direction is the same. They indicate the direction which a teacher should follow. The teacher can plan his class-room activities in consonance with the immediate and general objectives, which point in the same direction.

Sources of Objectives

Educational objectives are derived from different sources. Chief among these are: psychology of child development, social environment, economic structure, political order, cultural heritage, existing store of human knowledge, and the philosophy of life of a nation. Our society has undergone a tremendous change in the above mentioned factors during the last forty years and obviously the objectives must also change to suit the times. Objectives of education in U.S.A. will differ from those in the U.S.S.R. because of the difference in the philosophy of life of these two nations. Similarly the objectives of education in a developing country like ours will differ from those of a developed country like Japan.

Taxonomy of Educational Objectives

We have seen that an objective becomes meaningful when it is stated in terms of learning outcomes. These specific learning outcomes resulting from a course of study may run into hundreds. Most of them can however be classified under a relatively small number of headings. Any such classification is arbitrary but it provides a framework for classifying these outcomes and it directs attention toward changes in pupil behaviour in a variety of areas.

The taxonomy of educational objectives is basically a classification scheme just as the biological taxonomy is a classification scheme for animals into class, order, family, genus and species. Basically the taxonomy grew out of an attempt to resolve some of the confusion in communication which resulted from the translation of such terms as 'to understand', 'to appreciate', 'to think', etc. Meaning of these terms as implied by the curriculum-maker may differ when interpreted by the evaluator. Similarly teachers working in different institutions may interpret various terms in the syllabus in their own way if precise meaning of various terms is not communicated to them. Evaluation becomes faulty for want of communication between the curriculum-maker, the teacher and the evaluator. It was with this in mind that a group of college and university examiners under the leadership of Dr. Benjamin S. Bloom of the University of Chicago, attempted to devise a frame work or taxonomy that would give a precise meaning to various terms, provide a structure which would relate one term to another and provide additional meaning for a given term through this interrelationship.

A taxonomy of educational objectives* is a hierarchical classification of human growth in a given domain of development. Like biological taxonomies educational taxonomies provide classification of objectives which can be used in the field of education for various purposes. One of the most important purposes of a taxonomy is to establish accuracy of communication regarding instructional objectives and related matters. Education as a field of study is often criticized for having a large number of loosely defined terms and concepts such as scientific method, reasoning, appreciation, knowledge, understanding etc. A taxonomic framework of educational objectives should reduce this vagueness at least to a certain extent and become a means of more precise communication within the field of education. In the absence of a common understanding about hierarchical classification of

---

*Dr. R. H. Dave : *Taxonomy of Educational Objectives and Achievement Testing* [Published in 'Developments in Educational Testing' Vol. I. University of London Press,, London, 1969, pages 203-214.]

instructional objectives, the discussion of certain educational problems becomes nebulous and hence ineffective. Use of a taxonomy as a common scale for defining the boundaries of educational objectives should reduce vagueness and increase precision in professional discussions. A taxonomy can also become a means of understanding the sequence and organization of human development. A well developed taxonomy can be of great help in defining and evaluating educational standards of a school. The evaluator should know the relationship between taxonomies of educational objectives and various devices of evaluation.

For purposes of convenience the taxonomy was divided into three domains, the cognitive, affective and psychomotor. A child's development is always studied under these three categories. Handbook I, the *Cognitive Domain,* prepared under the leadership of Dr. Bloom was published in 1956. It deals with objectives having to do with thinking, knowing and problem solving. Taxonomy of Educational Objectives, Handbook II, the *Affective Domain,* was prepared by Krathwohl, Bloom and Masia and published in 1964. It includes objectives dealing with attitudes, values, interest, appreciation and social-emotional adjustment. The psychomotor domain covers objectives having to do with manual and motor skills. The feasibility of developing it is being studied by a group at the University of Illinois under Dr. Elizabeth Simpson.

Cognitive Domain*

Within this domain, the taxonomy tries to construct a comprehensive outline of the cognitive abilities that can be acquired from any body of subject matter. The following six major categories are listed in this domain :

1. Knowledge, 2. Comprehension, 3. Application,
4. Analysis, 5. Synthesis, 6. Evaluation.

(1) *Knowledge* : This is concerned with a pupil's ability to recall terms, facts, rules, principles etc. It is almost synonymous with the term information. Objectives

---

*Bloom, Benjamin S. and others : '*Taxonomy of Educational Objectives*' Handbook I : Cognitive Domain (1956).

that emphasize memorization of facts, principles, processes and patterns in different subjects fall under this category of the taxonomy.

The remaining five categories have been described as intellectual abilities and skills. They thus form one major class.

(2) *Comprehension* : It is defined as the lowest level of understanding. It is concerned with a pupil's ability to understand a given content to the extent that he can put it into his own words, summarize or explain it. The abilities that would be found in objectives in this category include being able to translate, give examples, illustrate, interpret, summarize or explain.

(3) *Application* : This is concerned with a pupil's ability to use rules, methods, procedures, principles and other types of generalization to new situations. The unfamiliarity and the problematic nature of a given situation evokes this level of thinking which is higher than mere comprehension. Objectives here would include being able to solve, predict, develop, explain or apply.

(4) *Analysis* : This category lays emphasis on the breakdown of the material into its constituent parts and detection of the relationships of the parts and of the way they are organized. The pupil would be able to depict the organization and the structure through which various elements are held together in a communication.

(5) *Synthesis* : It is concerned with a pupil's ability of putting together parts or elements to form a cogent whole. It involves uniqueness, originality and creative behaviour. This is reflected when one imagines the design of a machine. This is also reflected in the writings of a novelist or a playwright.

(6) *Evaluation* : This is concerned with judgments in terms of internal evidence or external criteria. It is defined as 'the making of judgments about value, purpose, ideas, methods, material etc.' In evaluating one has to appreciate the extent to which particulars are

accurate, effective, economical and satisfying. The judgments may be either quantitative or qualitative.

The taxonomy is hierarchical in nature, that is, each category is assumed to involve behaviour more complex and abstract than the previous category. Thus the categories are arranged from simple to complex behaviour and from concrete to abstract behaviour.

2. Affective Domain

Objectives dealing with attitudes, values, interest and appreciation are included in this domain. Meaning of these terms however is not uniformly understood by all. When we say that Sujata should appreciate art, do we mean that she should be aware of art work? Do we mean that she should seek it out and go to the museum? Should she experience an emotional kick or thrill when she sees any work of art? Should she be able to evaluate it?

This list could be extended but it is enough to suggest that the term 'appreciation' covers a wide variety of meanings and worse still, not all of these meanings are distinct from the terms 'attitude' and 'interest'. Thus if by appreciation we mean that a student should like a work of art well enough to seek it out, how do we distinguish such behaviour from *an interest* in art? If Sujata *values* art, does she have a favourable *attitude* toward art? Thus there is a confusion regarding the meaning of these terms. If the meaning of these terms is not clear how can the teacher achieve the objectives pertaining to interest, attitude and appreciation?

The committee of experts studying the affective domain tried to analyse the various objectives under this domain. The following characteristics may be pointed out:

(i) There is an emotional quality which is an important distinguishing feature of an affective response when we speak of interest, attitudes and appreciation.

(ii) There is an element of increasing automaticity as one progresses in the hierarchy.

(iii) An increasing willingness to attend to a stimulus is seen as one ascends the hierarchy.

(iv) The developing integration of a value pattern is seen at the upper levels of the hierarchy.

The authors used the concept of *internalization* as a common thread, summing through the developmental processes in the affective domain. Internalization refers to the inner growth that occurs as the individual becomes aware of and then adopts attitudes, principles, codes and sanctions which become inherent in forming value judgments and in guiding his conduct.

The taxonomy proposed by the committee* has five major categories :

(1) Receiving (Attending) (2) Responding (3) Valuing (4) Organization (5) Characterization by a value or value complex.

1. *Receiving or Attending* : It includes sensitivity to the existence of a certain stimulus and that the learner is willing to receive or attend to it. The first level in this category is awareness; the individual has his attention attracted to the stimuli. The second level is willingness to receive, and the next level is of controlled attention, that is, the individual controls his attention and he discriminates between various stimuli.

2. *Responding* : This is the category which many educators will find best described as the *interest* objective. At this level the learner becomes sufficiently involved in a subject or an activity. He gets satisfaction from the activity. At this level, the individual now differentiates the affective stimuli, he attaches emotional significance and value to them.

3. *Valuing* : This level describes increasing internalization as the person's behaviour is sufficiently consistent. There is the individual's commitment to a certain goal, idea or belief. The behaviour is quite consistent and stable. It involves not only acceptance of a value but preference for it and a commitment to or a conviction for a certain point of view.

---

*Kratwohl, Bloom and Masia : *Taxonomy of Educational Objectives* Handbook II—The Affective Domain (1964).

4. *Organization* : At this level, the learner builds up a value system that guides his behaviour. As the learner successfully internalizes the values, he encounters situations for which more than one value is relevant. This necessitates organizing the values into a system.
5. *Characterization by a value or a value complex* : At this level of internalization, the process reaches a point where the individual responds very consistently to value-laden situations with an interrelated set of values, a structure, a view of the world. The value-complex controls the behaviour of the individual. His behaviour is not controlled by stray emotions. At the highest stage of development the individual acts assiduously and consistently in according the value he has internalized. He attains an integration of his beliefs, and attitudes into a total philosophy or world view.

We can predict the behaviour of an individual if he has formed this value-complex. Will a person accept bribe in tempting situations? Will he be frustrated when he meets with a failure? We can answer such questions if the objectives in the affective domain are attended to and are being achieved.

3. **Psychomotor Domain**

This domain covers objectives having to do with manual and motor skills. A taxonomy for psychomotor domain is not yet available. Dr. Dave,* Head of the Department of Curriculum Development, NCERT, New Delhi, has proposed a taxonomy in the psychomotor demain, just in the form of a working hypothesis. The proposed taxonomy of educational objectives in this domain is based on the concept of co-ordination which runs as a common thread in the entire range of growth. Behaviours that belong to this domain include muscular action and require neuromuscular co-ordination. Our objectives in this domain would aim at development of proficiency in these acts by effecting co-ordination between various muscular actions performed by various

---

*Dr. R. H. Dave : *Taxonomy of Educational Objectivity and Achievement Testing* [Developments in Educational Testing, Vol. I, University of London Press, London 1969, Pages 203-214].

parts of the body. As this co-ordination increases, the learner's actions become more refined, quick and automatic.

Categories proposed in this domain are :

(1) Imitation (2) Manipulation (3) Precision (4) Articulation (5) Naturalization.

1. *Imitation* : When the learner is exposed to an observable action he begins to imitate that action. The performance, however lacks neuro-muscular co-ordination or control and hence is generally in a crude form.

2. *Manipulation* : This category emphasizes the development of skills in following directions. The learner is capable of performing an act according to instruction rather than just on the basis of observation as is the case at the level of imitation. He then begins to differentiate between various movements and is able to select the appropriate ones. He gradually moves towards fixation. The performance is fairly well set.

3. *Precision* : The proficiency of performance reaches a higher level of refinement in reproducing a given act. Here accuracy, proportion and exactness in performance become significant. The learner independently does the action. He shows confidence. He can control the speed as per requirements.

4. *Articulation* : This category emphasises the co-ordination of a series of acts by establishing appropriate sequence and achieving harmony among different acts. In driving for example, many acts are to be performed simultaneously by different parts of the body. The learner at this stage performs them in a harmonious manner.

5. *Naturalization* : At this stage, the skill attains its highest level of proficiency and the act is performed with the least expenditure of psychic energy. The act is carried out unconsciously; the habit of performance becomes the learner's *second nature*.

These five categories under the psychomotor domain need further discussion and thinking. They may therefore be considered purely as a tentative hypothesis.

Tripartite Structure

The development of a child takes place in three different ways, namely, cognitive, affective and psycho-motor. We can now build up a tripartite structure to obtain a comprehensive classification scheme of educational objectives. A diagram* showing all the three taxonomies is presented below :

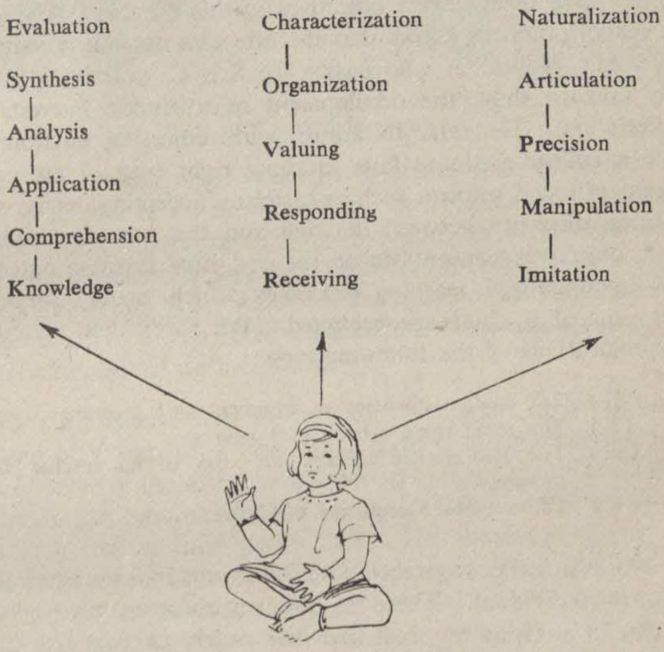


Fig. 2.4 : Exposing the child to various situations.
(A tripartite structure)

Uses of the Taxonomy

(1) The prime use of the taxonomy is the analysis and classification of objectives. The taxonomy, like the *periodic table* used in Chemistry, provides the panorama of objectives.

---

[*] Dr. R. H. Dave : *Ibid.*, p. 11.

(2) It is difficult to compare the school curricula of two or more States in India in the absence of the taxonomy framework. But the translation of objectives in the taxonomy framework can provide a basis for precise comparison of the two curricula. In a given subject we can develop a hierarchy of concepts by considering the hierarchy of objectives.

(3) The development of affective domain has pointed out the problems of achieving objectives in this domain. A study of the school practices shows that the affective domain is markedly neglected. A student who passes the S. S. C. examination generally fails to show the development of attitudes, interests and appreciation. He fails to adjust with changing environment. Even a college-graduate fails to show right type of attitude in matters of social reforms such as casteism, accepting dowry, social injustice, disparity between the rich and the poor etc. Instructional objectives concentrate on the cognitive learning outcomes. The emotionalized learning outcomes which mainly shape the behaviour of a child are neglected. We rarely find statements of specifications of the following type :

(i) The student discusses an argument *with pleasure*.
(ii) The pupil reads a book *with interest*.
(iii) The Headmaster assesses the work of the teacher *dispassionately*.
(iv) The student accepts the defeat *sportively*.

Words in italics from the above statements indicate some qualities of the individual. These are however unfortunately neglected. On careful analysis we shall find that nearly all cognitive objectives have an affective component, though it remains unspecified. We expect that a child should develop interest when it studies a given subject. This fact can be emphasized with the classification of objectives in the affective domain.

(4) In the cognitive domain we are concerned that the student *can do* a task when requested. In the affective domain we are concerned not with his ability but with his *will* or *desire* or *attitude* to do a certain task. The affective domain brings to light an extremely important and often missing element in cognitive objectives.

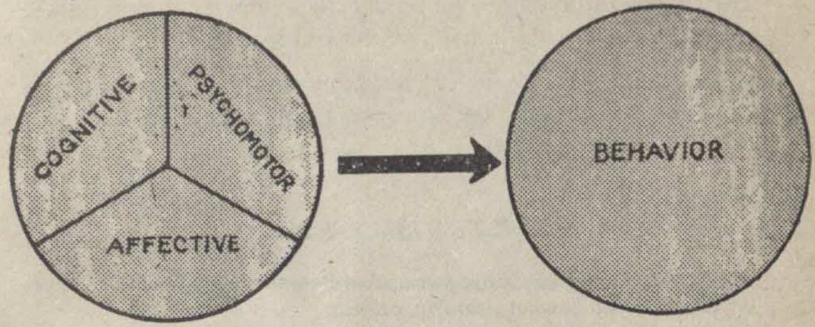


Fig. 2.5 : All round development.

(5) Shaping of one's behaviour depends upon the achievement of objectives of all the three domains (Fig. 2.5). Many of the schools over-emphasize the cognitive domain with the result that the development is lopsided. Just as we need a balanced diet for an all round development of the body, similarly balanced learning outcomes in the three domains are necessary for an all round development of an individual.

(6) The teacher tries to evaluate the change in behaviour at the end of a term or a year. But there is an extremely slow growth of some of the affective behaviours. It will be difficult to assess the formation of *an attitude* or the development of a *value system* in a student in a year's time. The measures of year's growth would reveal little change. This suggests that an evaluation plan covering at least a few years of school and involving the co-ordinated efforts of several teachers is probably a necessity. Such efforts would permit gathering longitudinal data on the same students so that gains in complex objectives would be measurable. Over a period of three or four years, the school should be able to assess the learning outcomes in the affective domain.

Because of the difficulty in measuring the changes in affective behaviours, a great deal of *erosion* takes place. As time passes on, the affective objectives cease to have any influence on the direction of the courses or in the choice of instructional activities. In part, this results from the fact that rarely are affective objectives reflected in the grading process. Students naturally concentrate on what counts in grading. We should, therefore, try to make these objectives more realistic parts of the courses.

With this background of the taxonomy of educational objectives, we shall consider the instructional objectives of school subjects in the next chapter.

FOR FURTHER STUDY

1. Explain the term objective and discuss how the instructional objectives influence the teaching-learning process.
2. Comment on the following statement put forth by a teacher: "My objective of teaching Geography is to teach the text-book prescribed for standad VIII."

   Explain how the above statement influences the teaching-learning process.
3. Why is it necessary to define an objective? How does it help in the process of evaluation?
4. (a) State with examples how instructional objectives are written.
   (b) Explain with suitable illustrations five characteristics for writing the specifications of an instructional objective.
5. Discuss the various sources of objectives and explain the relation between general and specific objectives or between ultimate and immediate objectives.
6. Explain the term taxonomy of educational objectives. How does the taxonomy help the educator and the evaluator?
7. State and explain the various terms of the taxonomy in the cognitive domain.
8. State and explain the various terms of the taxonomy in the affective domain.
9. Comment on the following statements:
   (a) Schools try to achieve objectives in the cognitive domain.
   (b) Objectives in the affective domain are generally neglected.
   (c) Emotional learning outcomes which shape the behaviour of a child are mostly neglected in the school.
   (d) Development of children appears to be lopsided during the school period.
10. Examine each of the following miscellaneous aims and indicate in what respects it is satisfactory or unsatisfactory as a statement of general objective:
    (i) To develop the power of appreciation and the ability to read. (Mother-tongue, Std. VII.)
    (ii) To develop competence in oral work and in problem solving. (Arithmetic, Std. V.)

(iii) To develop the ability to teach. [College of Education.]
(iv) To develop a better personality. [Standard IV.]
(v) To enable students to know Science. [Standard VIII.]
(vi) To study the prescribed text-book. [English, Standard X.]

11. Define the following statements of objectives in terms of learning outcomes. Give at least four learning outcomes :
    (i) To develop the ability to read with comprehension. [English, Standard VI.]
    (ii) To develop scientific attitude. [General Science, Standard X.]
    (iii) To develop skills. [Geography, Standard VIII.]
    (iv) To develop the power of understanding. [Mathematics, Standard XI]
    (v) To develop a sense of civic responsibility. [Civics, Std. IX.]

12. How will you modify the following statements of objectives for Standard V ?
    (i) To help pupils to understand concepts and formulae in mathematics and to enable them to apply the same.
    (ii) To develop skills and to develop interest in Science.
    (iii) To enable the students to speak and to write simple English.
    (iv) To velop in the pupil a national outlook and appreciation of the work of great men.

* * *

3

INSTRUCTIONAL OBJECTIVES

We have discussed in the previous chapter the taxonomy of educational objectives. This will help us in stating instructional objectives of various school subjects. The specific learning outcomes resulting from a course of study may run into hundreds; however we can classify them under a relatively small number of headings. The classification though arbitrary serves a number of useful purposes. It provides a framework for classifying the outcomes, and it indicates the type of learning outcomes that should be considered in any subject area. From the taxonomy we may select the following headings.

1. *Knowledge Objectives* : To acquaint the pupil with the terminology, specific facts, general principles, methods and procedures.

2. *Understandings* : Ability to generalize from given data, ability to recognize the limitations of data, ability to explain procedures and methods.

3. *Applications* : Ability to apply knowledge to novel situations, ability to interpret relationships.

4. *Skills* : These may vary from subject to subject such as : (i) laboratory skills, (ii) performance skills such as in art, music, sports etc., (iii) communication skills such as in conversation and in writing, (iv) computational skills, (v) work skills and (vi) study skills.

5. *Attitudes* : Social attitudes, scientific attitudes.

6. *Interests* : Personal interest, vocational interests.

7. *Appreciations* : Critical judgment of a thing appreciated, enjoyment of a thing appreciated.

8. *Adjustments* : Social adjustments, emotional adjustment.

[1] Instructional Objectives of English [Second Language]

In the teaching of English as a second language, emphasis would not be on the study of English literature, appreciation of poetry, study of various authors and their style. The pupil is expected to acquire necessary skills for communication. Hence emphasis is given on developing skills in reading, listening, conversing and writing. We shall consider some of the major objectives and translate them in terms of behaviour changes.

1. **Objective:** The student understands simple English when spoken.

Specifications in terms of behaviour changes:

1.1 He follows the pronunciation, accent and intonation of the speaker.
1.2 He understands the meanings of words.
1.3 He grasps simple ideas implied in the speech.
1.4 He follows the sequence of ideas expressed.
1.5 He recognizes relationships between objects, persons, ideas, events etc.
1.6 He follows directions given by the speakers.

2. **Objective:** The student speaks simple and correct English.

Specifications in terms of behaviour changes:

2.1 He uses proper stress and intonation.
2.2 He uses appropriate words and structures.
2.3 He uses appropriate pauses in speech.
2.4 He has a reasonable flow of speech.
2.5 He uses varied sentence patterns.
2.6 He presents his ideas in a proper sequence.

3. **Objective:** The student reads English with comprehension.

Specifications in terms of behaviour changes:

3.1 He locates key-words and phrases in a given passage.
3.2 He locates significant details.
3.3 He finds out the central idea.
3.4 He identifies relationship between main and subsidiary ideas

3.5 He grasps the meanings of words, phrases and sentences.
3.6 He locates the sequence of events, ideas and facts.
3.7 He draws inferences in respect of ideas, feelings, mood etc.
3.8 He discriminates between statements of facts and statements of views.
3.9 He attains reasonable speed and accuracy in reading.
3.10 He sifts relevant ideas from irrelevant ones.

4. **Objective :** The student writes simple English correctly.

Specifications in terms of behaviour changes :

4.1 He writes legibly and distinctly.
4.2 He uses appropriate words and phrases.
4.3 He uses varied sentence patterns.
4.4 He spells correctly.
4.5 He uses capital letters and marks of punctuation correctly.
4.6 He organizes ideas, facts, etc., into paragraphs.
4.7 He presents ideas and thoughts coherently and logically.
4.8 He makes proper use of sequence signals such as *but, besides, and, so,* etc.
4.9 He uses various forms in reporting.
4.10 He writes with reasonable speed.

5. **Objective :** The student acquires knowledge of the elements of the language.

Specifications in terms of behaviour changes :

5.1 The student

(i) recognizes, (ii) recalls, (iii) compares, (iv) discriminates, (v) identifies, (vi) classifies, (vii) locates errors, (viii) transforms, (ix) analyses, (x) synthesises, (xi) substitutes and (xii) illustrates in respect of grammatical structures, sentence-patterns, vocabulary, spelling, punctuation, pronunciation, etc.

5.2 He knows various forms of writing such as precis, letter, memorandum and essay.

The other learning outcomes of this objective would be realized through objectives 1, 2, 3 and 4 mentioned above and hence are not repeated.

6. **Objective :** The student develops interest in English.

Specifications in terms of behaviour changes :

6.1 He reads books in English other than the prescribed ones.
6.2 He reads the daily news in English.
6.3 He listens to radio broadcasts in English.
6.4 He writes articles, reports, stories, short poems for the school magazine.
6.5 He participates in debates, discussions and dramatics.
6.6 He recites poems.
6.7 He observes common English manners of speech.
6.8 He makes use of dictionaries and other reference books.

7. **Objective :** The student enjoys simple English literature.

Specifications in terms of behaviour changes :

7.1 He reads a literary piece of prose or poetry with due attention to stress, rhythm and emotion.
7.2 He enjoys the rhythm of the poem.
7.3 He shares the ideas and emotions of the author.
7.4 He quotes parallel lines from the mother-tongue.
7.5 He grasps the theme of the literary piece or the poem.
7.6 He recognizes word-pictures.

8. **Objective :** The student communicates effectively from English in his mother-tongue.

Specifications in terms of behaviour changes :

8.1 He understands the vocabulary and structures used in the passage as a whole.
8.2 He substitutes appropriate words, phrases, idioms and structures while translating a given passage in his mother-tongue.
8.3 He selects appropriate expressions to convey the sense of the passage.

8.4 He maintains the sequence of ideas as given in the original piece.
8.5 He maintains the spirit of the original passage.
8.6 He uses appropriate devices so that the translation reads as a connected whole.

It should be remembered that the specifications mentioned under each of the above objectives are only illustrative. One can further elaborate the list of specifications.

[2] **Mathematics**

Instructional objectives, when achieved, become learning outcomes. The expected learning outcomes are given below along with each of the instructional objectives. The list need not be considered as a final word. You can add some more learning outcomes under each head.

1. **Objective:** The pupil acquires knowledge of Mathematics.

Learning Outcomes : The pupil

1.1 recalls mathematical terms, processes, principles, formulae, definitions, etc.
1.2 recognises terms, instruments, processes, etc.

2. **Objective:** The pupil develops understanding in Mathematics.

Learning Outcomes : The pupil

2.1 gives illustrations for mathematical concepts, terms, principles, processes, etc.
2.2 detects errors in mathematical operations.
2.3 explains terms, concepts, principles, etc., in his own words.
2.4 observes mathematical relationships in various situations.
2.5 verbalises symbolic relationships and vice versa.
2.6 substitutes necessary numbers, symbols and formulae in mathematical operations.
2.7 verifies results.
2.8 compares mathematical concepts, processes, figures etc.

3. **Objective:** The pupil applies knowledge of mathematics to novel situations.

Learning Outcomes: The pupil

3.1 analyses a problem or data into component parts.
3.2 judges the adequacy, inadequacy or superfluity of data.
3.3 establishes relationships among the data.
3.4 gives a number of methods of solving a problem.
3.5 selects the most appropriate formulae or principle or method or process to solve a problem.
3.6 reasons deductively.
3.7 reasons inductively.
3.8 makes generalizations.
3.9 draws inferences.
3.10 predicts results on the basis of data.

4. **Objective:** The pupil acquires skills in:

(A) *Handling mathematical instruments with ease.*

Learning Outcomes: The pupil

4.1 draws fairly satisfactory free-hand figures.
4.2 selects the most appropriate mathematical instruments.
4.3 takes necessary precautions in taking measurements of drawings.
4.4 takes measurements correctly.

(B) *Drawing geometrical figures and graphs.*

Learning Outcomes: The pupil

4.5 draws figures to given specifications.
4.6 draws figures quickly.
4.7 uses appropriate markings to denote different parts of a figure.

(C) *Computation.*

Learning Outcomes: The pupil

4.8 does oral calculations correctly.
4.9 does oral calculations quickly.
4.10 does written calculations correctly.
4.11 does written calculations quickly.

4.12 uses correct notations and symbols.
4.13 avoids unnecessary steps in the solution of a problem.
4.14 is systematic in the working of a problem.

(D) *Reading of tables, charts, graphs, etc.*

Learning Outcomes : The pupil

4.15 selects appropriate mathematical tables.
4.16 uses mathematical tables, charts, ready reckoners, etc., correctly.
4.17 co-ordinates the different sections of the table correctly.
4.18 reads graphs correctly.

5. **Objective :** The pupil develops interest in Mathematics.

Learning Outcomes : The pupil

5.1 reads literature on mathematics from books, magazines, journals and newspapers.
5.2 reads biographies of mathematicians.
5.3 discusses mathematical problems with other students.
5.4 takes pleasure in solving problems.
5.5 sticks to the problem and does not leave it unsolved.
5.6 constructs mathematical problems and puzzles.
5.7 solves mathematical problems and puzzles on his own.
5.8 participates in mathematical club activities.
5.9 prepares models, charts, graphs and diagrams illustrating mathematical principles.
5.10 collects material of mathematical interest such as typical problems, puzzles, biographies, models, etc.

6. **Objective :** The pupil develops positive attitude towards mathematics.

Learning Outcomes : The pupil

6.1 encourages activities related to mathematics (as listed under interest).
6.2 admires good students, teachers of mathematics and mathematical literature.
6.3 is ready to undertake tests related to mathematics even at the cost of personal inconvenience.
6.4 does not avoid persons, objects, literature related to mathematics.

7. **Objective:** The pupil develops scientific attitude through the study of mathematics

Learning Outcomes : The pupil

7.1 examines various aspects of a solution, before arriving at a conclusion.
7.2 does not accept any proposition unless it is logically established.
7.3 seeks fresh clues beyond the ones already known.
7.4 reports observations without cooking up the results.
7.5 tries to understand the other person's point of view.
7.6 accepts his mistakes without reservations.
7.7 keeps an open mind and does not regard any argument as final.

8. **Objective:** The pupil develops appreciation for mathematics.

Learning Outcomes : The pupil

8.1 recognises the qualities of exactness, compactness, brevity, precision, accuracy and logical coherence in mathematics.
8.2 visualises the beauty of mathematical relationships such as similarity, symmetry, etc., in specific life situations.
8.3 explains mathematical principles involved in simple machines, models, etc.
8.4 suggests improvements in figures, models and simple machines, in keeping with mathematical principles.
8.5 creates beautiful or pleasing patterns with the help of geometrical figures.

9. **Objective:** The pupil develops personality traits such as concentration, punctuality, accuracy, neatness and perseverance.

Learning Outcomes : The pupil

9.1 is not disturbed by extraneous stimuli while solving mathematical problems.
9.2 listens to arguments attentively.
9.3 sticks to a task until it is completed.
9.4 does not get confused by failures.
9.5 readily undertakes new tasks.

9.6 puts forth arguments when convinced about their correctness without any hesitation.
9.7 points out errors in argument or working without hesitation.

[3] Geography

1. **Objective:** The pupil acquires the knowledge of facts, terms, concepts etc. in Geography. (acquisition of knowledge)

Learning Outcomes : The achievement of this objective is seen when the pupil

1.1 recalls facts, terms, concepts, principles, processes, etc.
1.2 recognizes facts, terms etc.

2. **Objective:** The pupil develops understanding in Geography.

Learning Outcomes : The change in behaviour as seen through the achievement of this objective would be :

2.1 He discriminates between different facts, terms etc.
2.2 He detects errors in statements.
2.3 He compares and contrasts various statements.
2.4 He classifies various terms and concepts.
2.5 He identifies relationship between cause and effect.
2.6 He cites illustrations.
2.7 He interprets data presented in various forms.
2.8 He intrapolates simple gaps in data.

3. **Objective:** The pupil applies knowledge of Geography to novel situations. (application of knowledge).

Learning Outcomes : The pupil who has acquired this ability :

3.1 Identifies and analyses the problem.
3.2 Selects relevant facts, principles etc., to explain a new situation.
3.3 Judges adequacy of data or any other evidence.
3.4 Reorganizes the material in a new manner.
3.5 Establishes relationship among the data.
3.6 Draws inferences from given information.
3.7 Predicts outcomes in a given situation.

3.8 Sets up cogent hypothesis to explain geographical facts, phenomena etc., in a novel way.

4. **Objective**: The pupil acquires skills related to Geography.

Learning Outcomes : The pupil who has acquired these skills :

4.1 Draws maps, charts and diagrams to depict geographical information.
4.2 Reads charts, maps, diagrams etc.
4.3 Detects errors of accuracy and proportion in a given diagram and corrects them.
4.4 Handles tools, apparatus, etc.
4.5 Presents data in symbolic form.
4.6 Preserves specimen and other exhibits of geographical interest.
4.7 Makes observations accurately.

5. **Objective**: The pupil develops interest in Geography.

Learning Outcomes : The pupil who has developed interest :

5.1 Reads on his own literature related to Geography.
5.2 Reads biographies and stories of explorers and geographers.
5.3 Observes geographical phenomena keenly.
5.4 Visits places of geographical interest.
5.5 Participates in the activities of the school geography-club.
5.6 Pursues useful hobbies such as collection of specimens of rocks, crops, pictures showing the life of people of various lands.

6. **Objective**: The pupil develops positive attitude through the study of Geography.

Learning Outcomes : The pupil who has developed this attitude :

6.1 Understands human life as influenced by physical environment and develops understanding and sympathy for other groups of society.
6.2 Recognizes the significance of interaction between man and his environment.

6.3 Critically examines problems.
6.4 Formulates judgments free from prejudices.
6.5 Realizes the importance of conservation of natural resources.

7.. **Objective:** The pupil develops desirable personality traits.

Learning Outcomes : Following traits are reflected in the behaviour of the pupil :

7.1 Inquisitiveness.
7.2 Tolerance for other peoples' ways of life.
7.3 Appreciation of the contributions made by explorers and travellers.
7.4 Perseverance.

The list of learning outcomes given above is just illustrative and not exhaustive. Teachers may add more on the basis of their experience.

[4] Mother Tongue

The emphasis in the teaching of the mother-tongue will certainly differ from that in the teaching of the second language. In the latter case, emphasis will be on acquiring necessary skills of communication such as reading, writing and conversation. In the teaching of the mother-tongue, however, these skills are acquired in the elementary school stage; emphasis in high school classes would naturally be on the study of the language, appreciation of poetry, study of various authors and their style, and development of desirable personality traits.

Various objectives of teaching the mother-tongue are achieved mainly through oral work, reading and written work.

1. **Objective:** To develop the power of oral expression.

Learning Outcomes : The achievement of this objective is judged from the various skills acquired by the pupil in oral work. Specifications of these skills in terms of behaviour changes are given below :

1.1 He distinguishes correctly between various sounds and tones.

1.2 He follows the pronunciation, accent and intonation of the speaker.
1.3 He understands the meaning of words.
1.4 He grasps ideas implied in a speech.
1.5 He follows the sequence of ideas expressed in a speech.
1.6 He uses appropriate words and structures in his speech.
1.7 He expresses his ideas effectively and coherently.
1.8 He expresses proper feelings through his speech.
1.9 He speaks with fluency, grace and politeness.
* 1.10 He expresses effectively through discussion, debate and dialogue.
* 1.11 He speaks before an audience with ease and confidence.
* 1.12 He uses a progressively larger vocabulary.
* 1.13 He has a reasonable speed and flow in speech.
* 1.14 He speaks in complete sentences.

2. **Objective:** To develop the ability to read.

Reading and Recitation : In the lower classes, emphasis will be given on oral reading. Silent reading should be introduced from standard III.

Oral Reading

2.1 He reads aloud clearly with correct pronunciation.
2.2 He reads with proper accent and emphasis.
2.3 He reads with understanding.
2.4 He reads with fluency and punctuation.
2.5 He recites songs and poems in chorus and individually with appropriate action, rhythm and recitation.
2.6 He expresses proper feelings through proper intonation

Silent Reading

2.7 He locates keywords and phrases in a given passage.
2.8 He locates significant details.
2.9 He finds out the central idea.

---

* Learning outcomes marked with an asterisk are to be emphasized in higher classes.

* 2.10 He identifies relationship between main and subsidiary ideas.

2.11 He discriminates between statements of facts and statements of views.

* 2.12 He draws inferences in respect of ideas, feelings, mood etc., of various characters depicted by the author.

* 2.13 He compares the style of various authors.

* 2.14 He reviews a book.

2.15 He attains reasonable speed in reading.

* 2.16 He recognizes idioms and proverbs occurring in a passage.

3. **Objective:** To develop the ability to write.

Learning Outcomes : In the lower classes emphasis will be on writing a good legible hand while in higher classes it will be on using the language for self-expression.

3.1 He writes legibly and distinctly.

3.2 He uses appropriate words and phrases.

3.3 He uses varied sentence patterns.

3.4 He organizes ideas, facts etc., into paragraphs.

* 3.5 He uses various forms of writing such as essay, story, dialogue, etc.

* 3.6 He takes notes while listening a lecture.

* 3.7 He presents the ideas and thoughts coherently and logically.

3.8 He writes with reasonable speed.

3.9 He observes rules of orthography in writing.

* 3.10 He uses figures of speech, idioms, proverbs in his writing.

* 3.11 He writes a review of a play, a novel, a story etc.

* 3.12 He expresses his feelings through his writing.

---

* Learning outcomes marked with an asterisk are to be emphasized in higher classes.

4. **Objective:** To acquire knowledge of the elements of the language.

Learning Outcomes : The pupil is expected to know grammar, punctuation and rules of orthography.

4.1 He recognizes sentence as a unit of thought.
4.2 He recognizes the parts of speech.
4.3 He uses appropriate parts of speech in sentences.
4.4 He recalls rules of orthography.
4.5 He locates orthographical errors and corrects them.
4.6 He writes to dictation.
4.7 He identifies various kinds of sentences.
4.8 He transforms a sentence of a given pattern into another.
4.9 He recognizes the derivation of words.
4.10 He analyses compound words.
4.11 He transforms a word by using prefixes and suffixes.
4.12 He identifies figures of speech, idioms, and proverbs in a passage.

5. **Objective:** To develop interest in the study of the mother-tongue.

Learning Outcomes : The pupil who has acquired interest in the study of mother-tongue will behave differently from the one who has not acquired such interest.

5.1 He likes reading books other than the prescribed ones.
5.2 He recites poems.
5.3 He participates in discussion.
5.4 He reviews fiction, plays, stories etc.
5.5 He writes articles, short poems etc.
5.6 He participates in the activities of the literary club.

6. **Objective:** To develop the power of appreciation.

Learning Outcomes :

6.1 He reads a literary piece of prose or poetry with due attention to stress, rhythm and emotion.
6.2 He locates key-words and ideas.
6.3 He shares the emotions of the author.

6.4 He identifies figures of speech and quotes parallel instances.
6.5 He quotes various illustrations in his writing.
6.6 He dramatizes and expresses emotions.

7. **Objective:** To acquire desirable personality traits.

Learning Outcomes : This aspect of acquiring personality traits is not very important in the study of the second language but it cannot be neglected in the study of the mother-tongue. Many desirable qualities have to be inculcated in the pupils. The study of the mother-tongue is the most powerful instrument for this purpose. People in India have become conscious in political matters but they do not appear to be so in matters of social reforms. We pay lip-service in matters of social-injustice, untouchability, caste-system, accepting dowry, disparity between the rich and poor etc. The pupil must show a change in his behaviour in matters of social changes mentioned above. This is possible if the teacher emphasizes this aspect in her teaching. How will the pupil who has developed those qualities behave ?

7.1 He accepts social reforms.
7.2 He practices the social reforms.
7.3 He shows sympathy to the victims of social injustice.
7.4 He sacrifices his interests for the benefit of those who have been denied justice.
7.5 He respects the good in the society.

These learning outcomes cannot be achieved in a year or so. But the change in behaviour should appear over a period of 4 or 5 years of schooling.

[5] General Science

In the teaching of this subject, our main objective is to acquaint the pupils with the basic principles of science and to develop in them the scientific attitude. It is not intended that schools should produce scientists. Objectives can be classified as follows.

1. **Objective:** The pupil acquires knowledge of the fundamentals of Science. [knowledge objective]

Learning Outcomes : The pupil who has acquired such knowledge,

1.1 recognizes certain terms, facts, principles, objects, specimens, instruments etc.
1.2 recalls various terms, principles, facts etc.

2. **Objective :** The pupil develops understanding in science.

Learning Outcomes : The pupil who has developed understanding,

2.1 uses symbols for words or words for symbols.
2.2 uses appropriate units to express various physical quantities.
2.3 classifies substances, terms etc., into given groups.
2.4 illustrates terms, concepts, processes etc., by suitable examples.
2.5 detects errors or defects in certain instruments, experimental procedures, definitions etc.
2.6 rectifies errors in experimental procedures.
2.7 makes certain calculations based on certain scientific relationships.
2.8 recognizes proper sequence in a given process.
2.9 verifies a rule, a principle or a property.
2.10 judges the adequacy or superfluity or data.
2.11 connects related phenomena into meaningful integrated units.
2.12 distinguishes between facts and opinions.

3. **Objective :** The pupil applies knowledge of science to novel situations. [application objective]

Learning Outcomes : The pupil who has acquired this ability,

3.1 explains scientific phenomena by giving reasons.
3.2 makes predictions in a given situation.
3.3 verifies the predictions.
3.4 draws inferences and makes generalizations.
3.5 suggests procedures for experiments.
3.6 establishes cause and effect relationship.
3 7 devises new procedures or modifies the old ones.
3.8 suggests improvisation of apparatus or experiments.

4. **Objective :** The pupil develops laboratory skills.

Learning Outcomes : These skills can be observed through the process of performance as well as through the product.

(a) **The process of performance :** The pupil

4.1 plans the arrangement of apparatus.
4.2 checks various instruments and rectifies errors if they are within his competence.
4.3 measures material accurately.
4.4 reads scientific instruments carefully.
4.5 performs experiments with neatness.
4.6 takes precautions in handling instruments.
4.7 makes accurate observations of parts, specimens, processes etc.
4.8 records observations and calculates where necessary.
4.9 adapts himself with new instruments and novel experiments.

(b) **The product of performance :** The pupil

4.10 draws diagrams, graphs etc.
4.11 calculates and finalizes the results.
4.12 interprets data and draws conclusions.
4.13 explains orally the procedures, principles etc., involved in the experiment.
4.14 records experimental procedures and conclusions.
4.15 tabulates correctly.
4.16 improvises home-made apparatus.
4.17 dismantles and cleans the apparatus.

5. **Objective :** The pupil develops interest in science.

Learning Outcomes : The pupil who has developed interest

5.1 reads scientific literature.
5.2 collects specimens, models, pictures of scientific interest and preserves them.
5.3 seeks pleasure and delight in preparing models, tools etc.
5.4 actively participates in science-fairs, science-exhibitions, science-clubs, field-trips etc.

5.5 seeks opportunity in observing and studying natural and man-made surroundings.
5.6 selects science projects and works upon them.

6. **Objective :** The pupil develops scientific attitude.

Learning Outcomes : How does a pupil who has developed this attitude differ from the one who has not developed this attitude ? This attitude is reflected in such qualities as open-mindedness, intellectual honesty and suspended judgment. The pupil

6.1 suspends all judgment until all the relevant facts are available.
6.2 seeks cause and effect relationship in the data.
6.3 is clear and precise in his statements and activities.
6.4 makes interpretations which are free from bias.
6.5 shows willingness to consider new interpretation and thus to revise the old ones.
6.6 does not accept opinions of others merely on authority, but tries to seek the validity of the argument.
6.7 does not believe in superstitions.
6.8 indicates confidence in the data obtained by scientific procedure.
6.9 is prepared to review his own arguments and opinions.

7. **Objective :** The pupil develops a sense of appreciation.

Learning Outcomes : The pupil who has developed this sense,

7.1 realizes the importance of science in human progress.
7.2 derives a sense of pleasure in understanding the advance of science and technology in the modern world such as interplanetary travel, use of atomic energy etc.
7.3 takes delight in understanding the contribution of science in day-to-day life such as advertizing, synthetic products etc.,
7.4 appreciates the contributions made by scientists to human progress.
7.5 expresses joy, excitement and thrill when he knows the achievements of scientists.

8. **Objective :** The pupil develops desirable personality traits.

Learning Outcomes : The pupil who has acquired these traits,

8.1 shows inquisitiveness.
8.2 is clear and precise in his arguments.
8.3 shows tolerance for the view point of others.
8.4 shows perseverance.
8.5 insists on accuracy and neatness.

[6] Social Studies

The social studies deal primarily with past and current problems of human relationships and with the interactions of human beings as they associate with one another in varied political, economic and social activities. Such school subjects as history, geography, civics, sociology, economics and problems of democracy are included in this area. Objectives of this subject can be classified as follows.

1. **Objective :** To acquire the knowledge of facts, terms, concepts etc., in social studies.

Learning Outcomes : The pupil who has acquired such knowledge,

1.1 recognizes certain terms, facts, principles, concepts etc.
1.2 recalls various terms, concepts, principles, facts etc.

2. **Objective :** To develop understandings in social studies.

Learning Outcomes : The pupil

2.1 realizes the role of existing human institutions as they have evolved through the past and their inter-relationship.
2.2 understands the importance of democratic way of Government and its utility in human welfare and happiness.
2.3 identifies the problems of local community and the need for wide participation in community concerns by all citizens.
2.4 realizes the need for co-operative living and interdependence among groups and nations.

2.5 recognizes the historical and geographical forces underlying the behaviour of groups and nations.
2.6 identifies relationship between cause and effect in social data.

3. **Objective :** To develop the ability to apply knowledge of social studies to novel situations.

Learning Outcomes : The pupil

3.1 identifies and analyses significant aspects in social problems.
3.2 distinguishes facts from opinions.
3.3 recognizes bias, prejudice and other distortions in social situations.
3.4 discriminates between issues in terms of their relevance to particular social problems.
3.5 draws inferences from given social situations.
3.6 predicts outcomes in a given situation.

4. **Objective :** To acquire necessary skills.

Learning Outcomes : The pupil

4.1 takes part in group discussion.
4.2 takes part in co-operative planning.
4.3 shows ability to think reflectively on social problems.
4.4 evaluates ideas and opinions on controversial problems expressed through radio, movies, newspapers etc.
4.5 presents data in different forms through maps, charts, graphs etc.
4.6 interprets social data from tables, charts, graphs etc.

5. **Objective :** To develop right type of attitude.

Learning Outcomes : The pupil who has developed this attitude will certainly differ from the one who has not developed this attitude. He

5.1 shows faith in the democratic way of life.

5.2 shows respect for the dignity and worth of individual personality irrespective of his caste, creed, race, colour etc.
5.3 believes that every individual is entitled to equal rights to life, liberty and the pursuit of happiness.
5.4 is willing to sacrifice his own interest if they clash with those of the society.
5.5 suspends judgment on social problems until the information is adequate to permit conclusions.
5.6 revises his conclusions if additional reliable information is obtained.
5.7 shows recognition that for every privilege we enjoy, there is an accompanying responsibility to self and to society.
5.8 shows respect for the law.
5.9 believes that social problems creating tension in the society can be solved through democratic processes.

6. **Objective:** To develop interest in social studies.

Learning Outcomes: The pupil

6.1 participates in various school activities.
6.2 shows awareness in one's rights and duties.
6.3 discusses problems related to civic life.
6.4 keenly observes the behaviour of an individual as seen in a group.

7. **Objective:** To develop desirable personality traits.

Learning Outcomes: If the various objectives mentioned above are achieved, there will be a corresponding change in the behaviour of the pupil, which will indirectly lead to the development of desirable personality traits. Such a pupil

7.1 is willing to sacrifice his own interests if they clash with those of the society.
7.2 tries to adjust himself when he stays in a group.
7.3 shows tolerance for other people's views and ways of life.
7.4 accepts the verdict given by the majority even if his personal views differ from those of the majority.

7.5 is sensitive to problems such as poverty through disparity, exploitation of the weaker section etc.
7.6 does not use unfair means for achieving the ends in social life (e.g. accepting or giving bribe, using black money, adulteration of food etc.)
7.7 does not misuse his authority in tempting situations.

[7] Work Experience

The objective of introducing work-experience in schools is mainly to develop interest and to enable the pupils to acquire manipulative skills. The school is not expected to prepare experts and technicians through this activity. Let us consider wood-work as one of the subjects under this category. Not much emphasis will be laid on 'knowledge' objective. It is a skill subject and hence emphasis should be laid on acquiring necessary skills.

1. **Knowledge Objective:** To acquire the knowledge of the elements of wood-work.

Learning Outcomes: The pupil

1.1 recognizes various technical terms.
1.2 recognizes various tools used in wood-work and their parts.
1.3 recalls various terms, concepts, principles etc.

2. **Understanding Objective:** To develop understanding in the theory of wood-work.

Learning Outcomes: The pupil

2.1 calculates the cost of finished products.
2.2 reads a blue print.
2.3 keeps a record regarding the speed and output of work.
2.4 locates any defects in the tools.

3. **Skills Objective:** To develop necessary skills.

Learning Outcomes: The pupil who has acquired necessary skill will do the work with necessary speed and accuracy.

The pupil

3.1 handles such tools as a chisel, a saw, a plane etc.
3.2 carries out minor repairs.
3.3 prepares 'working drawings' of the articles, showing the plans and elevations with measurements.
3.4 does exercises in sawing, planing and chiselling.
3.5 prepares jointless articles like handles, pegs, rulers, pointers etc.
3.6 does gluing, finishing and polishing.
3.7 applies lacquer and paints to the articles prepared.
3.8 prepares articles involving joints. like a blotter, a tray, a lamp-stand, book-shelf etc.
3.9 sharpens the tools.
3.10 shows the desired speed in doing the work.
3.11 prepares the model with accuracy.

4. **Interest Objective :** To develop interest in wood-work activities.

Learning Outcomes : The pupil

4.1 accepts this activity as a hobby.
4.2 spends his leisure-time in wood-work activities.
4.3 participates in school exhibition and displays his articles.
4.4 collects material that is useful for gaining efficiency.

5. **Attitude Objective :** To develop a sense of dignity of labour.

Learning Outcomes : The pupil

5.1 values the dignity of labour.
5.2 participates in activities which involve labour.
5.3 considers that the labourer also has a worthy place in the society.
5.4 does not look-down upon activities involving labour.

6. **Objective :** To develop desirable personality traits.

Learning Outcomes : Practice in work experience should lead to the inculcation of desirable traits. The pupil

6.1 regards self-help as an important characteristic.
6.2 insists on accuracy and efficiency.

6.3 is not satisfied with mediocre type of work.
6.4 uses his leisure-time advantageously.

[8] Physical Education

About thirty years ago this subject was not given a place in the school-curriculum. If the ultimate objective of school education (Chapter II, page 25) is the all-round development of child's personality, we have to include all subjects that will contribute to the achievement of the ultimate objective. Accordingly subjects like work-experience, physical education, fine-arts are given a place in the present day curriculum. We must also know the limitations of these subjects. Neither do we intend to prepare all of our students for Olympic games nor do we want them to work as acrobats. The curriculum will vastly differ if we have these objectives in view. The teacher will therefore have to view the programme of physical education in right perspective.

Objectives :

1. To develop a good physique, organic fitness and neuromuscular skills.
2. To build up habits of health, correct postures and movements.
3. To awaken an interest in wholesome play and recreation.
4. To give satisfactory outlet for healthy expression of the child's spontaneous activities.
5. To develop qualities such as smartness, alertness, initiative enterprise resourcefulness, self-discipline, team-spirit, leadership, sportsmanship and love for service leading to good citizenship.

Learning Outcomes : Various objectives given above can be specified in terms of behaviour-changes on the same lines as has been done so far in this chapter. It will be seen that emphasis is not laid on the *knowledge* objective but it is laid on the development of *interest, attitude* and *skills* which will ultimately lead to the development of desirable personality traits.

[9] Art

It is now a recognized fact that a child has a natural desire to express himself in some form. It is because of this desire for expression that a child feels delight in creative activities. A child's initial efforts at these activities though rough and crude according to adult standards, are full of life and vigour. The aim of art education in schools is not to produce artists but to foster and develop this creative self-expression innate in every child, through various media like chalk, charcoal, colours, clay, crape-paper etc. Art in schools is generally interpreted as *drawing* and *painting* only. However at the early stage of the child's development no cleavage between art and craft is desirable. The child should be free to select whichever material and method he is inclined to adopt. In the art syllabus for lower primary schools, therefore, activities in art and craft work should be taken.

Objectives :

1. To train the eye in the observation and discrimination of forms and colours.
2. To develop the capacities of visualization and imagination.
3. To develop memory for forms.
4. To develop the capacity for tasteful design and decoration.
5. To enable the child to acquire increasing control over tools and material.
6. To release the creative impulses and suppressed emotions of the child.
7. To develop the child's aesthetic sense and cultivate in him a taste for an appreciation of the beautiful in nature and art.
8. To develop desirable personality traits in the child.

Learning Outcomes : Various objectives given above can be specified on the same lines as has been done so far. It will be seen that emphasis is to be given on objectives pertaining to skills, interest and appreciation which will lead to the development of the child's personality.

[10] Music

In modern days music plays an important part in the social life of an individual. Music helps in developing an individual's aesthetic taste and in satisfying his aesthetic sense. It thus contributes to the enrichment of his life.

Objectives :

1. To create and cultivate a taste for music.
2. To enable the pupil to enjoy music in life after school-days.
3. To train the pupil to hear and appreciate good music.
4. To encourage those who reveal special aptitude for developing *executive skills,* vocal or instrumental.
5. To provide opportunities for corporate activities based upon music.

Learning Outcomes : Teacher should not lay emphasis on teaching the grammar of music. Otherwise children may dislike the subject. She should allow children to recite poems in chorus so that they enjoy singing of poems and songs.

The objectives may be specified in terms of pupil-behaviour. Emphasis should be on the achievement of skill, appreciation and interest.

FOR FURTHER STUDY

1. State the objectives of teaching the mother tongue. Specify these objectives in terms of learning outcomes.
2. Bring out clearly the difference in the learning outcomes of teaching the mother tongue to standards IV and X.
3. 'The aim of art education in schools is not to produce artists'. Comment on this statement and discuss the objectives of art education in schools.
4. Specify the following objectives in terms of learning outcomes :
   - (a) To acquire computational skill in mathematics. (Std. VII)
   - (b) To develop the power of oral expression through the mother tongue. (Std. IV)
   - (c) To develop interest in science. (Std. X)
   - (d) To develop right type of attitude through the teaching of social studies. (Std. IX)
   - (e) To develop the child's aesthetic sense through art. (Std. VI)
   - (f) To develop the ability to write simple English. (Std. VIII)
   - (g) To cultivate a taste for music. (Std. VII)
   - (h) To develop desirable personality traits through the teaching of the mother tongue. (Std. X)
5. *Comment on the following statements* :
   - (a) A teacher says to her class, "Let us try to develop interest in science during this period."

(b) A teacher preaches, "Let us try to develop our personality through the study of lesson XI in the text-book of the mother tongue."

(c) A teacher says, "My pupils of Std. I have no aesthetic sense. The pictures they draw are very rough and crude."

(d) A teacher tries to teach the grammar of music to children of Standard IV.

(e) A student studies civics but fails to show the qualities of a good citizen.

(f) A college graduate takes into consideration the caste of the candidate while voting at the time of elections though he has studied civics in his college.

(g) A student of standard VIII writes correctly the description of a thermometer but he cannot read a thermometer.

6. State the objectives of teaching the following subjects and specify them in terms of expected behaviour changes:

(i) English as a second language (std. VIII).
(ii) English as the monther tongue (std. VIII).
(iii) Mathematics (std VI).
(iv) Geography (std. IV).

* * *

4

OBJECTIVES OF TEACHER EDUCATION

We have so far discussed instructional objectives of school subjects. We shall now consider the objectives of the Teacher Education Programme. There are many critics who say that teacher education programme as practised to-day is a failure and it needs to be radically improved. How much truth is there in the above view point ? For a critical study of the programme, we must start with the objectives of Teacher Education. What type of teacher do we want to turn out from a college of education ? What competencies should we expect of such a teacher ? What type of attitudes should she display ? Which type of teacher is liked by the pupils ? These are our reference points, while deciding the objectives.

The student teacher should not only know the technique of class-room teaching, but she should also be acquainted with the daily school-routine. She should also know why and how children learn, and under what conditions they might learn even more effectively. Hence while framing the curriculum of a teacher-training course, we must include subjects which will acquaint the students with the Foundations of Education and also with the Practice of Education. The course content is thus conditioned to the objectives of teacher-education. We may therefore first consider the objectives.

Objectives of Teacher-Education Programme

1. To help the student teacher to understand the aims of education.
2. To promote in her an awareness of the role of the school and of the teacher in realising these aims and ideals.
3. To develop an understanding of the close relationship between the school and the society.
4. To help her develop competence to teach subjects of her specialization.

5. To help her understand the teaching-learning process and the factors affecting the process.
6. To enable her to gain insights about herself and her pupils.
7. To enable her to feel and appreciate the dynamism inherent in the teaching profession.
8. To motivate her through the realization of this dynamism to become an effective teacher.
9. To develop in her interest for the profession.
10. To foster a positive attitude towards the teaching profession.

The above objectives can be classified under the usual categories such as knowledge, application, skills, interest, attitude etc. We have to frame the curriculum so as to fulfil the above objectives. We can now classify the objectives and give their specifications. The list of learning outcomes given under each head is not exhaustive. Additional items can further be added.

1. **Knowledge Objective :** To help the student-teacher to acquire the knowledge of subjects necessary for her professional preparation.
    - (i) Foundations of Education (Philosophical and Sociological).
    - (ii) Foundations of Education (Psychological).
    - (iii) School organization and Educational Administration.
    - (iv) Methods of Instruction and Evaluation Procedures.
    - (v) Methodology of School Subjects.
    - (vi) Problems in the field of Education.

Learning Outcomes should be stated with reference to the subjects mentioned above.

2. **Application Objective :** To help the student-teacher to apply the principles of education in various situations in the school.

Learning Outcomes : The student-teacher

2.1 carries out experiments to study various mental processes.

2.2 observes various pupils and records individual differences.
2.3 maintains various records required in the school.
2.4 administers tests and interprets the results.
2.5 motivates pupils.
2.6 carries out surveys to study the influence of environment on the development of a child.

This list can further be made exhaustive by referring to various activities of a training college.

3. **Skills Objective:** To help the student-teacher acquire necessary skills in teaching.

Learning Outcomes : The training college will have to organize practice teaching in such a way that the student-refer to Chapter II and Chapter XIII.]

3.1 introduces the topic methodically.
3.2 creates readiness among the pupils.
3.3 encourages students to participate in the discussion.
3.4 asks thought-provoking questions.
3.5 inspires pupils for further study.
3.6 controls the class efficiently.

[This list can be made exhaustive. For additional items refer to Chapter II. and Chapter XIII.]

4. **Interest Objective:** To develop interest in the teaching profession.

Learning Outcomes : The student teacher who has developed interest in teaching will behave differently from the one who has not. Such a teacher

4.1 reads educational literature.
4.2 shows keenness in understanding children's behaviour.
4.3 plans her teaching.
4.4 looks out for new ideas in making teaching effective.
4.5 enjoys teaching in the class-room.
4.6 gives a patient hearing to students.
4.7 studies problems of student-behaviour.

5. **Attitude Objective:** To develop a positive attitude towards the profession.

Learning Outcomes : The student teacher develops a sense of professional responsibility to the school, and to the society. The training college will have to consider the attitude of the student-teacher while evaluating her. Such a teacher

5.1 is conscious of her duties.
5.2 respects individuality of the child.
5.3 observes professional ethics.
5.4 has a critical outlook about her work.
5.5 is willing to consider new ideas. .
5.6 is proud of the teaching profession.
5.7 believes that the teacher can mould the personality of the child.
5.8 believes that the child must be the centre of various activities.
5.9 believes in the improvement of teaching methods through research.
5.10 believes that example is better than precept.

6. **Appreciation Objective:** To develop the power of appreciation.

Learning Outcomes : The student-teacher

6.1 recognizes the contributions of great educators.
6.2 realizes the importance of education in the life of an individual.
6.3 appreciates the behaviour patterns of children on account of individual differences.

7. **Objective:** To develop desirable personality traits among the student-teachers.

Learning Outcomes : The personality of the teacher is an important factor in the process of education. In a training college attention should be paid to the development of desirable personality traits. If this aspect is neglected, it is most likely that the student-teacher will fail miserably in doing her job in later life. Some of the desirable traits which pupils appreciate are given here. A teacher should try to match her characteristics with those given below :

7(A) **A teacher liked by pupils —**

7.1 is helpful to pupils.
7.2 is cheerful, happy and has a sense of humour.
7.3 explains the subject matter clearly.
7.4 is interested in pupils and tries to understand pupils.
7.5 makes work interesting, creates desire to work.
7.6 commands respect.
7.7 is impartial, shows no favouritism.
7.8 is not sarcastic.
7.9 has a pleasing personality.
7.10 is patient, kindly and sympathetic.
7.11 is fair in marking and grading the pupils.
7.12 has intellectual honesty and moral integrity.

7(B) **How does a teacher, not liked by pupils behave ? Such a teacher —**

7.1 is not helpful, does not explain lessons well.
7.2 does not plan her work.
7.3 does not smile; loses temper easily.
7.4 indulges in favouritism.
7.5 is mean, unreasonable, intolerant and ill-mannered.
7.6 is unfair in marking and grading.
7.7 is not interested in pupils; does not understand them.
7.8 is too loose in discipline, has no control over the class.
7.9 brings in too many irrelevant personal matters.
7.10 talks too much about herself.
7.11 cannot set any ideal before the pupils.
7.12 only preaches but does not act accordingly.
7.13 has no pleasing personality.

Retrospect : How do we evaluate a student-teacher ? We lay emphasis only on the knowledge and skills objective. Other objectives are neglected with the result that the student-teacher does not bother much to achieve the remaining objectives. The student-teacher leaves the college with a wrong impression that the training programme is meant for learning a few skills and knowing a few secrets of the trade. Training colleges will not be condemned if all the objectives are attended to and evaluation is done accordingly. We have not to produce acrobats in a training college, by teaching a few skills.

The course content in a training college will have to be determined on the basis of objectives discussed above. Decide the programme according to the type of teacher you want in the school. There must be an experimental school where the student-teachers can carry out some experiments. There must be a workshop-cum-hobby centre where the trainees can prepare various teaching aids. Trainees should get an opportunity for participating in various school activities and also for observing the pupils in various situations. All the programmes of a training college must be geared to the various objectives. Evaluation of the student-teachers must be done with reference to the statements of objectives.

The training of teachers has to be fundamentally different from training in other professions. We must realise that a teacher in a school is first and foremost *a teacher* and only secondarily *a technician* for transmitting certain knowledge and skills to the young minds effectively. If her main role is that of a technician, our emphasis in a training college would be on initiating the trainees into the arts and skills of class-room teaching. But the teacher is more than a technician. She must know the developing mind. She has to inspire the children. She must have the dynamism necessary for moulding the life of the pupils. Hence the training college programme should help the teacher to play this role effectively. How can this be done? We shall discuss this aspect in the next chapter.

FOR FURTHER STUDY

1. State the objectives of Teacher Education at the secondary school level.
2. What qualities would you expect of a teacher from a pre-primary school?
3. Specify the following objectives in terms of learning outcomes:
   (i) The teacher should be able to apply the knowledge of psychology to the class-room situation.

(ii) The teacher should have an insight about herself and her pupils.
(iii) The teacher should have competence to handle a class.
(iv) The teacher should have a positive attitude towards the teaching profession.

4. Which of the objectives of teacher-education are neglected in the present day training colleges? Suggest remedies for improving the situation.
5. Which objectives are more important at the pre-primary stage and which ones at the secondary school stage?

* * *

5

LEARNING EXPERIENCES

We discussed instructional objectives of various subjects in the previous chapters. They determine the ends or goals of education. We have now to consider the means to attain these goals. Learning occurs when the pupil has an experience, that is, when he reacts to the situation in which he finds himself. Learning takes place by interaction between the situation and the learner. When any situation acts upon the learner, the learner reacts, he modifies his behaviour and this interaction results in learning. Consider the following examples :

(i) Four year-old Dileep burnt his fingers when he touched a hot pan on the stove. There is some interaction. He does not now go near the stove. Learning has taken place through this experience.

(ii) Seven year-old Swati sees a person slipping over a banana skin on the street. She now does not step over a skin fallen on the street. She throws the skin in a dust-bin. This experience has modified her behaviour.

(iii) Children from Koyna-nagar school experienced the tremours of an earthquake in 1967. They learnt how to face a calamity. This experience helped them modify their behaviour.

(iv) Madhao, a student from standard X observed that his teacher was careless in his work and hence his services were terminated. He learnt that careless behaviour is not tolerated in the society.

(v) Sampat, a six-year old boy from the slum area observes that elderly persons abuse one another even with a slight instigation in everyday life. He reacts to the situation and picks up the same style. He has learnt though not in a desired direction.

(iv) Sujata participated in the school tournaments. She was cheered-up by all when she won a prize. The experience modified her behaviour. She developed confidence and she pursued these activities further.

(vii) Four year-old Vidya tries to pour water from a glass into a bottle. She gains control over her muscles. Many such activities are carried out in a Montessori school. All these help in modifying the behaviour (Fig. 5.5).

It will be seen from these varied examples that a child is exposed to various situations in life. Every experience as if impinges on the child who in turn reacts. This interaction results in learning which is seen through modification of the behaviour.

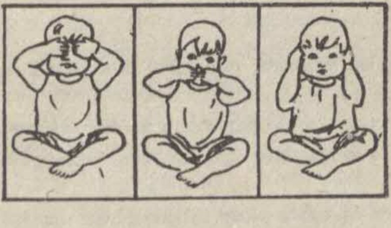


Fig. 5.1 : No learning experiences.

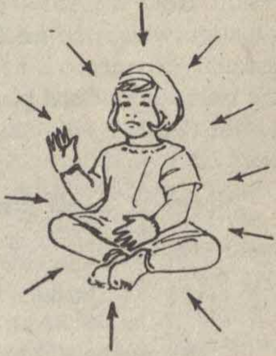


Fig. 5.2 : A child exposed to learning experiences.

Experiencing is simply seeing, hearing, feeling, tasting, smelling and so on. The individual reacts to these experiences and learning takes place. We learn because things happen to us and we do something in turn. It is only through experience that we learn. Learning will not take place in the absence of any experiences (Fig. 5.1). The phrase *learning experience* is probably a tautology. All experiences are learning experiences.

Every moment the child is learning because of varied experiences he is getting in life. The kind of learning which takes place is a result of the kind of experience which the child gets. In the traditional school environment what type of experiences are provided ? The teacher gives factual information and the child memorizes the information by rote. The learning process becomes synonymous with memorization and factual information becomes the learning product. But learning does not mean only memorization. It is defined as the process of acquiring knowledge, abilities, skills, attitudes and so on. These are the learning outcomes. What kind of experiences should be provided to produce these outcomes? We cannot allow the child to meet with stray experiences in the school. They have to be planned. *Quality of experiences* which the children have, will determine the *quality of learning.*

Various Aspects of a Learning Experience

Children learn by thinking, feeling and doing. Learning results from the active participation of the pupil to the stimulus situation which the teacher creates in the class. Remember that learning experience is not a part of a syllabus, nor is it a unit nor the traditional lesson plan nor mere activity. It is the interaction of the learner with the situation created by the teacher, for example:

(i) A teacher praises those students who have completed their home-work.

(ii) A teacher encourages students to participate in the discussion.

(iii) Two students abuse one another; this has been disapproved by the teacher.

(iv) Students by turn have to tell a story to the whole class.

(v) Students listen to the news-bulletin of A.I.R. and then they write the important news-item in their note-books.

In all these situations, there is an interaction between the learner and the situation. This then becomes a learning experience. Each of these situations helps to modify the behaviour of the pupil.

Classification of Experiences

Learning experiences can be classified as direct and indirect or vicarious experiences.

Direct Experience : This refers to learning activity that involves first hand experience with various objects or symbols. It includes what is commonly referred to as perceptual learning, but it is a broader term since it includes experiences with symbols such as used in mathematics or science. Perceptual learning applies to experiences dependent upon seeing, hearing, tasting, smelling, feeling, handling and manipulating things in different ways. When we use such terms as sweet-bitter, soft-hard, tall-short, we get the meaning through perceptual learning. We can then use these symbolic words for describing various things. Following are illustrations of direct experiences :

(i) Observing and experimenting with materials and apparatus.
(ii) Constructing models, plans etc.
(iii) Drawing, painting, dramatization.
(iv) Presenting a description either orally or in writing.
(v) Summarizing, stating generalizations etc.
(vi) Listing important facts, points and so on.

Vicarious or Indirect Experience : This is the type of experience through which we acquire the learning outcome without first hand experience. It makes use of direct experience of others. We learn much through the experience of others. These experiences include such activities as reading, looking at pictures, listening to lectures and discussions and so forth. Such learning activities are very important since it is impossible for us to have all first hand experiences. You can enjoy the thrill of mountaineering by listening to a talk of a mountaineer. The best way to understand what a revolution is, would be to live through one and have direct experience; but since the teacher is prohibited from organizing revolutions even for educational purposes, she does the next best thing. She does it through vivid description using books, pictures, movies and other material. We must keep in mind constantly however that the ability of the learner to profit by the experiences of others depends upon the extent and adequacy of his previous first hand or direct experience. Following are illustrations of vicarious experiences :

(i) Reading accounts or discussions in books, magazines, papers etc.
(ii) Listening to oral discussions, lectures and so on.
(iii) Observing pictures, maps, charts, models etc.

It is obvious that in practical learning situations, we cannot completely separate the activities according to these two types, nor is it desirable to do so. There will be a combination of direct and vicarious experiences in many of the activities.

Criteria for choosing a Learning Experience

(i) It should be directly related to the instructional objective.
(ii) It should be meaningful and satisfying to the learner. Playing with clay would be meaningful and satisfying to children going to pre-primary schools but it may not be so for adolescents.
(iii) It should be related to life-situations.
(iv) It should be appropriate to the maturity level of the learner (Fig. 5.3).

Fig. 5.3: A child is not a small model of a man.

(v) While selecting the experience, availability of material and of time should be considered.
(vi) It should be varied and rich in content. A pupil who observes the habits of birds and their nests has a richer experience than the one who merely sees the habits of one bird.

While selecting any learning experience, the teacher must see that it leads to the attainment of any given educational objective. She should remember that it is possible to have a variety of learning experiences for the attainment of a single objective.

It is also possible that one learning experience may lead to the attainment of several objectives. Changes in pupil's knowledge, attitude and skills have to be brought about gradually. No single experience will generally bring about the total change. Many experiences, one reinforcing another will have to be provided. They may also have to be repeated in increasing complexity or levels, in meaningful sequence extended over a period of time. Some of these may be for the whole class, others for different groups and a few may be for the individual. Learning experiences are not merely confined to the class-room. They can be developed outside also and should compliment each other.

Various Centres of Learning Experiences

1. *Home* : It is one of the influential centres of education. The child learns a number of things at home. Some children get enough material for reading and for games. The parents influence the child's behaviour through their own actions.
2. *Society* : This is another agency which informally provides a number of learning experiences. These may not necessarily bring about a change in the desired direction. Lottery, gambling are some of the illustrations which may have an adverse effect while some other experiences have a wholesome effect. People who work in the society with a sense of dedication provide desirable learning experiences.
3. *Mass-Media* : Radio, TV, newspapers, libraries enrich the lives of individuals by providing rich and varied material.
4. *Play-Ground* : It is our experience that many qualities are developed on the play-ground. Skill in various games, attitude of co-operation, leadership-qualities are likely to be developed through activities on the play-ground.
5. *Journeys and Excursions* : Value of excursions has been recognized since long. When you go for example from Bombay to Mysore, you get such a varied experience which

influences your behaviour. You become more confident, you face the difficulties boldly. You see people of various types and languages.

6. *Peer-Groups* : A child learns a number of things while he is among his peers. Qualities such as co-operation, perseverance, honesty, helpfulness are developed in the peer-group (Fig. 5.4).

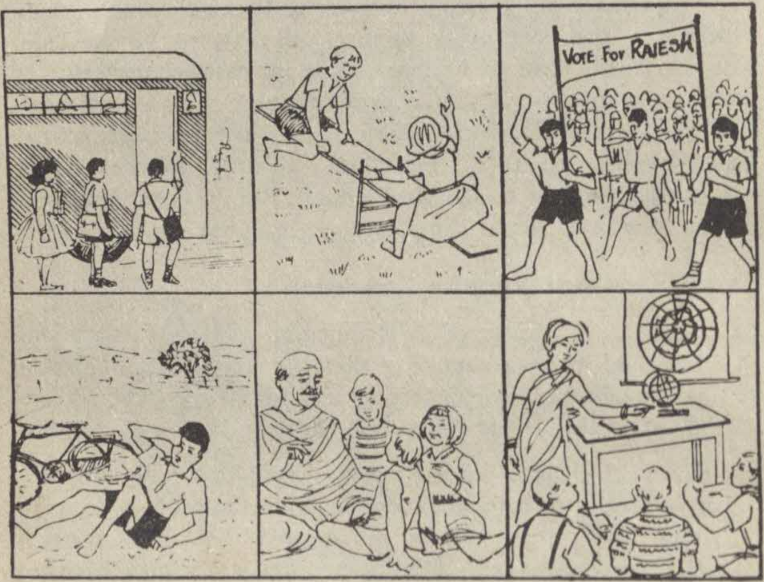


Fig. 5.4 : Learning Experiences.

The teacher may not have control over all the above centres which provide learning experiences. She will have therefore to use various teaching devices for providing suitable learning experiences. Following are some of these :

(i) *Experiments* : They provide learning experiences to achieve objectives such as skill in handling apparatus, skill in observation, understanding.

(ii) *Audio-Visual Aids* : Radio, films, filmstrips, epidiascope etc., enrich teaching and develop pupil interest. They provide learning experiences which widen the scope of understanding and develop interest in the subject.

(iii) *Co-curricular Activities* : Debating, dramatization, competitions etc., provide learning experiences which help in achieving objectives pertaining to the development of expression.

(iv) *Devices of Teaching* : Various devices of teaching are useful for providing learning experiences.

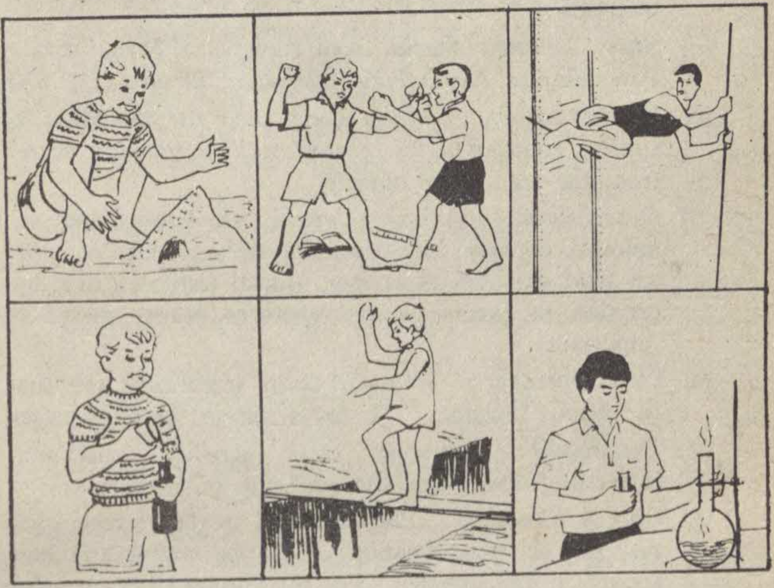


Fig. 5.5 : Learning Experiences.

We shall now consider learning experiences pertaining to various school subjects. The list given below is not exhaustive. A teacher can add various other activities, but the activity should be objective-oriented.

[1] Mother Tongue

Main objectives of teaching the mother tongue are developing (i) the power of oral expression (ii) the ability to read and (iii) the ability to write. Learning experiences will have to be geared mainly to these objectives.

(1) *Oral Expression* : Experiences in oral work should be organized through the following activities; some of these activities are suitable for elementary school children while some are suitable for middle school classes.

(a) *Conversation* : (i) Through such topics as self, home and surroundings, occupations of children's relatives, news items, festivals in the community etc. (ii) Polite conversation on social occasions (iii) Participation in an interview.

(b) *Story Telling* : Stories from Ramayana, Mahabharata, Panchatantra, Aesop, Fairy tales and Folk tales (Fig. 5.4).

(c) *Dramatization* : Action songs and dramatization of stories, enacting a few dialogues and dramatic scenes from the reader and outside.

(d) *Description* : (i) Asking the pupils to describe domestic animals, common birds, familiar trees, pictures etc. (ii) seasonal activities of farmers, social festivals (iii) description of excursions and visits to nearby places of importance.

(e) *Short speeches* : To make short speeches at the time of various functions. To recite selected prose passages.

(f) *Discussion* : Planning class activities, discussing interesting news items and problems of the class.

(g) *Speech Training* : Phonetic drill in the correct, clear and distinct pronunciation of all the vowels and consonants in the alphabet. For the purpose of ear training and speech training, it is essential that throughout the primary stage, the teacher shows the pupils, by example, how to produce the right voice and tone, how to form various sounds, how to pronounce difficult words, and how to read a passage with appropriate pauses, stress, emphasis, speed and tone.

(2) *Reading and Recitation* : Following activities will provide varied learning experiences.

(a) Model reading by the teacher, oral reading by individual students with due attention to proper accents, emphasis and natural intonation.

- (b) Giving passages for silent reading and testing for comprehension.
- (c) Supplementary reading material should be provided through story books. Pupils should discuss what they have read.
- (d) Recitation of prose passages and also of poems from the reader.
- (e) Recitation of poems and songs from other books.
- (f) Organizing competitions in recitation, in story-telling etc.
- (g) Reading of newspapers followed by discussion of news items daily for some time.
- (h) Developing vocabulary through spelling games, giving jumbled words, giving words beginning or ending with certain letters, giving synonyms and antonyms etc.
- (i) Encouraging students to use a dictionary.

(3) *Written Work* : Following activities will be found useful in providing suitable learning experiences.

- (a) Transcription on the slate in bold and legible letters of simple sentences from the black-board.
- (b) Writing to dictation words and simple sentences.
- (c) Writing about five simple sentences on the topics covered in oral work.
- (d) Practice of writing between ruled lines.
- (e) Copy writing and writing to dictation short paragraphs.
- (f) Developing a story from given points.
- (g) Describing in good language pupil's own experiences about any incident or activity.
- (h) Writing of letters.
- (i) Writing answers to simple questions from an unseen passage or a poem.
- (j) Using words, phrases, idioms, figures of speech in one's writing.
- (k) Writing to dictation using various marks of punctuations.
- (l) Developing vocabulary through word-building and word-games, giving antonyms and synonyms, compound-words.
- (m) Competitions in hand-writing and in essay-writing.
- (n) Preparing a news-bulletin of the class.

(o) Organizing literary clubs, collection of good thoughts, poems, songs, story pictures, anecdotes etc.

The above activities should be organized after taking into consideration the maturity level of the child. Some of these activities are suitable for pupils of standards I to IV while remaining ones are suitable for pupils of higher standards.

[2] Mathematics

Learning experiences can be provided through various activities such as counting, measuring, oral drill, etc.

1. Counting objects, reading and writing numbers.
2. Recognition of coins.
3. Giving practice in building 100 primary facts of addition.
4. Giving practice in understanding 100 primary subtraction facts.
5. Recognition of common shapes, comparing sizes of objects.
6. Practice in handling money.
7. Using multiplication tables.
8. Giving practice for improving the speed of doing the basic skills.
9. Carrying transactions in a post-office, a bank, a shopping-centre etc.
10. Using metric measures of length, weight, area, volume etc.
11. Preparing bills and writing simple accounts etc.
12. Reading the time, reading a calendar etc.
13. Drawing various geometrical figures, bisecting an angle, drawing a triangle or any other figure as per given conditions.
14. Problem solving : Students should be given a challenge through problems. Ready-made solutions should not be given.
15. Developing a problem : Students should be asked to prepare examples of their own, based on the theorem or the principle already taught. This is a very thrilling experience.
16. Checks and cross-checks : Students should be trained in applying various checks to the solution of examples

In the lower classes emphasis should be laid on activities through concrete situations. In higher classes emphasis should be on reasoning.

[3] History

1. Effective narration by the teacher leads the students as if to live through the historical incident.
2. Dramatization : Students get acquainted with people's way of life in the past.
3. Dramatization is useful for inculcating such qualities as devotion for work, sacrifice for the nation, perseverance in getting across the hurdles in one's way etc.
4. Visits to places of historical importance : These places as if narrate the history more effectively.
5. Collection of various specimens having historical importance.
6. Visits to museums.

[4] Civics

1. Practising the civic sense in the school, for example, standing in a queue when the situation demands, using the dust-bin for throwing waste material, keeping the school premises clean etc.
2. Conducting discussions in the class-room : One can express his viewpoint, he must also listen to another person's viewpoint; only one at a time should speak. These are some of the habits that can be developed in the school.
3. Debates, seminars and symposia should be organized to show that a subject can be studied by considering various aspects. We may not like another's viewpoint but we must allow him to express his views.
4. Holding elections for students' council. Let the candidates get an opportunity to propagate their candidature. Let them prepare an election manifesto. This will inculcate values essential for democratic life (Fig. 5.4).
5. Let the students see the working of the Zilla-Parishad, Municipality or the Corporation. Let them attend the

meeting and see how the Councillors discuss various matters.

6. Let the students see a village or a town and study its problems such as facilities for drinking water, sanitation, communication etc.
7. Chorus singing of the National Anthem.
8. Citizenship training through the organization of camps.

[5] Geography

1. Using the globe of the earth to study the rotation of the earth, and the phenomenon of day and night.
2. Use of relief-maps for explaining such concepts as valleys, plateaus, mountains, slopes etc.
3. Reading of maps (Fig. 5.4).
4. Drawing the outlines and filling them up.
5. Recording daily temperature, pressure, rain-fall etc.
6. Introducing the various geographical regions through pictures and movies.
7. Talks to be given by those who have visited various countries.
8. Excursions for studying climatic conditions at various altitudes.
9. Visits to fields, factories and markets for studying such factors as raw material, means of communication, demand for finished products, climatic conditions of the place etc.
10. Observation of the sky to study the planets, formation of eclipses, and various phases of the moon.

[6] General Science

The teacher can use narration as an effective learning experience in a subject like history; however she will have to use demonstration, experimentation, observation for giving necessary learning experiences in the teaching of science. The list given here is only illustrative.

1. Demonstrating experiments in the class-room and encouraging students to observe and derive conclusions.
2. Allowing students to perform experiments themselves.
3. Allowing students to handle apparatus, to locate the defects if any and set it in order (Fig. 5.5).
4. Developing a museum with the help of students.
5. Developing a garden in the school premises.
6. Encouraging students to maintain records of growth of plants under varying conditions, through day-to-day observation.
7. Maintaining records of daily temperature, pressure etc.
8. Creating problem solving situations.
9. Encouraging students to find out cause and effect relationship in day-to-day situations, developing scientific attitude through discussion and experimentation.
10. Observing the stars and planets at night.
11. Visit to a planetorium.
12. Conducting a science-club in the school.
13. Observing habits of birds and animals.
14. Preservation of food.

[7] **English** (Second Language)

1. Using spelling games.
2. Reading of story-books from the library; asking the students to narrate the story in the class-room, and writing it down in the note-books.
3. Practice in copy-writing for improving hand-writing.
4. Encouraging students to use the dictionary.
5. Recitation of poems and of selected passages.
6. Oral reading: Daily practice in oral reading should be carried out.
7. Practice in conversation in the class-room.
8. Listening to the news-bulletin of AIR.
9. Writing of essays, letters, precis etc.
10. Conducting dummy interviews for acquainting the students with its technique.
11. Practice of writing answers to questions based on the text and unseen passages.

[8] Physical Education

Major objective is to develop skills in children. Learning experiences must be geared to this objective.

1. Developing various physical skills such as running, jumping, hopping, rope-climbing and various exercises in athletics (Fig. 5.5).
2. Developing skill in games : Practising the various games daily.
3. Conducting tournaments and athletic meets.
4. Organizing camps : These are useful for inculcating qualities such as tolerance, co-operation, adjustment, leadership etc.
5. Practice in swimming, playing foot-ball or cricket, skating, climbing a steep rock etc.
6. Developmental exercises including free movements.
7. Rhythms and action songs.
8. Recreative games.
9. Drill and marching.
10. Folk dances.

[9] Work-Experience

Let us consider only one subject namely carpentry.

1. Using various tools : Practice to be given in sawing, planing, chiselling etc. (Fig. 2.2).
2. Preparing articles without joints, such as pegs, rulers, pointers etc.
3. Preparing articles with joints such as a blotter, a tray, a book-shelf, a lamp-stand etc.
4. Conducting a hobby workshop.
5. Practice for achieving skill and speed.
6. Minor repairs of various tools such as a chisel, a saw, a plane, a hand-drill etc.
7. Gluing up, finishing and polishing the articles.

[10] Art

1. Making patterns on floor, table, slate etc., with objects like sticks, beads, brass-rings, flowers etc.

2. Drawing designs on floor, slate etc. with chalk or crayons.
3. Using coloured rangoli, sawdust etc. for decoration.
4. Preparing simple decorative designs based on forms derived from familiar objects such as fruits, leaves, flowers, birds, fish etc.
5. Clay modelling: animals, birds, fruits, vegetables etc.
6. Pictorial and imaginative drawing and painting.

The teacher should give various learning experiences in art by encouraging the child in his creative attempt by giving him full scope to express his ideas and experiences, imagination and fantasy freely, unhampered by the imposition of adult view point of good art.

[11] Music

1. Practice in chorus singing of various devotional, folk and marching songs.
2. Practice in singing the National Anthem.
3. Providing opportunities for vocal music to individual pupils who show aptitude for it.
4. Providing opportunities for instrumental music.
5. Listening various records of musicians.
6. Organizing a school orchestra.

[12] Teacher Education

Practice teaching is a major activity in training colleges. Student teachers are asked to give a few stray lessons in some classes. This learning experience cannot be considered as rich and varied.

Student teachers must be acquainted with various problems which the teacher has to tackle in her day-to-day work. Hence the student teachers must work in the practising or co-operating school for a period of about two weeks where she can get the following experiences:

1. Teaching practice: Teaching a class one or two subjects continuously for 2 weeks (Fig. 5.3).
2. Observation of pupils: Noting individual differences, studying some individual cases such as a shy student, a very bright student, a problem child etc.

3. Assessing the home-work of students.
4. Maintaining cumulative records.
5. Discussion with parents through the parent-teacher association.
6. Preparing a diagnostic test and administering it.
7. Using remedial measures.
8. Writing report cards to parents.
9. Discussion with individual students for locating their problems.
10. Guidance to students.

Unfortunately only the first item of this list is practised in many of the colleges. Other items are neglected. All the above learning experiences will help the student-teachers in achieving the objectives of practice-teaching. Remember that practice-teaching is not meant only for teaching a few tricks of the trade to the student-teacher. Teacher is not an acrobat working in a circus. Hence the idea of showing only a few tricks of class-room teaching must be discarded.

Limitations

It is very difficult to give direct experience in all matters. Experiences such as tremors of an earthquake, pangs of a woman at the time of delivery, feelings of a girl on her wedding day, Armstrong's feelings when he landed on the Moon on July 21, 1969 etc., have to be described. Every student cannot get a direct experience in such matters. Poets and novelists have described these experiences beautifully in their writings. A student can appreciate them through imagination.

Emotionalized Learning Experiences

If we study the learning experiences provided in schools, we find that the objectives of the cognitive domain are being achieved but the affective domain is almost neglected with the result that the child *knows* a certain subject but *he may not be interested in it*. In shaping one's behaviour, the affective domain is important. Too often the teacher thinks of *learning* simply as memorizing facts, getting knowledge from a book or acquiring skill by

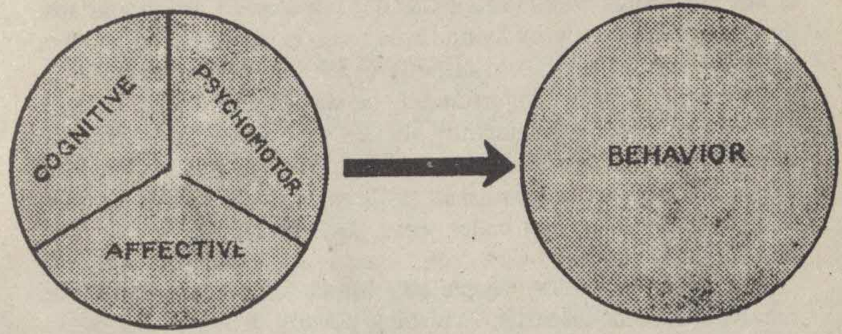


Fig. 5.6: Learning experiences shaping the behaviour.

practice of some kind. Her attitude reduces learning to a mechanical process. She teaches a lesson, she uses some teaching aids, she gives some homework, all of these being done in a routine way. She pays little attention to the development of right type of attitude and of interest among the children. The child studies science without developing the scientific attitude. Many educated persons criticize the evil practices in our society such as accepting dowry, practising untouchability etc., but they themselves adhere to these practices. The truth is well expressed in the old adage, "*Knowledge is no guarantee of right conduct.*"

Teachers should therefore emphasize learning experiences of the affective domain in their work (Fig. 5.6). Ultimately these learning experiences are going to shape the behaviour of the child. We should remember that our own welfare and the welfare of the society in general depend quite largely upon the emotionalized learning products. If a student securing first-class in the final examination fails to develop such qualities as self-control, civic-sense, sympathy for the poor, disgust for injustice, respect for other's rights he will not be a useful member of the society. He may turn out to be an *intelligent rascal.* He may even turn out to be a misfit.

The teaching of attitudes, appreciations, interests and habits of good behaviour has been on the whole, unsatisfactory in the school. We have failed to give students the kind of experience from which such emotionalized responses may be expected. Let it be emphasized again that explaining a thing is not equivalent to the thing itself. The desired emotional response is aroused only by living through an experience that touches the thing itself. A devoted disciple of Christ was working for the betterment of lepers. Unfortunately he was affected by the disease. His friends expressed sympathy for his unfortunate lot. But the disciple was not worried. He said, " When I was serving the lepers I could not imagine their distress; now I can live through that experience and can better serve the lepers. "

We find that many people pay lip service in such matters as removing income-disparity, banishing poverty, doing social justice to the down-trodden etc. But when it comes to making a sacrifice of one's own interests, these very people begin to hesitate. This happens only because this important aspect of formation of attitudes which controls an individual's behaviour, has been neglected in their education.

In providing learning experiences of the affective domain, teacher's personality is a dominant factor. If the teachers lacks enthusiasm, if she lacks the attitude she seeks to create, you cannot expect desired outcomes such as enthusiasm and formation of attitudes among the students. Enthusiasm begets enthusiasm. Inspiration is catching and only the active, enthusiastic teacher can succeed well in this work. The teacher must embody the ideals, attitudes, and habits of behaviour she seeks to instil into the lives of the students. Example is a powerful influence and it is always better than precept. Students soon detect insincerity and sham on the part of teachers. They are fatal to any prospect of effective teaching.

In India the progress of social reconstruction is going on with a snail's pace, because school education has been emphasizing the cognitive aspect only. A student becomes only literate when he passes the secondary school-certificate examination but he is not educated in the real sense of the term. The formation of attitudes, appreciations, interests, ideals and habits of behaviour is being neglected in school education. Teachers should therefore see that suitable learning experiences are provided in the school for developing these qualities.

FOR FURTHER STUDY

1. Explain the term learning experience. Illustrate how a child gets these experiences even in out of school situations.
2. State the various characteristics of a good learning experience.
3. Discuss the learning experiences that you will plan for achieving the following objectives :
   (i) Developing students' vocabulary of English as a second-language in standards V, VI and VII.
   (ii) Ability to express one's ideas orally. (Mother tongue).
   (iii) Ability to read silently with comprehension. (Mother tongue).
   (iv) Developing interest in reading books. (Mother tongue).
4. Give a classification of learning experiences and illustrate each type with suitable examples.
5. State the importance of emotionalized learning experiences. How do they help in shaping the behaviour of a child ?
6. Which type of learning experiences should be planned in the following situations ?
   (i) To develop a democratic way of life among students.
   (ii) To develop workmanship in carpentry.
   (iii) To appreciate paintings.
   (iv) To develop interest in science.
7. Consider the following units from the syllabii of standard VII. (i) Practical Geometry (ii) Plant life (iii) Geography of Ceylon (iv) Simple composition in English. Select any two of the above units and prepare a list of suitable learning experiences.
8. Choose any two units from the following and prepare a list of suitable learning experiences mentioning the class in view.
   (i) Graphs.
   (ii) Digestive system of animals.
   (iii) Rights and duties of citizens.
   (iv) Climate of India.
   (v) Story-writing (mother-tongue).
9. Discuss a few learning experiences that would be useful in developing the personality of a child.
10. Discuss why the present day school-practices are inadequate in shaping a pupil's behaviour.
11 Suggest ten learning experiences for student-teachers so that they acquire necessary skill in teaching.

* * *

6

THE UNIT PLAN

A teacher cannot successfully teach any subject unless she is able to plan her teaching skilfully. A well-planned lesson will not in itself ensure success since it is difficult to achieve various objectives in one lesson. Because good planning is basic to good teaching, the beginning teacher owes it to himself to become thoroughly familiar with planning procedures. Since the teacher has a major responsibility in planning the learning activities in the courses she teaches, much of the effectiveness of her teaching will depend upon the efficiency of her planning.

Too frequently, the teacher thinks of planning a course in terms of a text-book. There are however many other important factors to be considered. One of the problems the teacher should understand is the selection of objectives and subject matter. She has also to consider the learner. In planning the scope of a course, the teacher should be familiar with various plans such as topic plan, unit plan, project plan, contract plan etc. At present the unit plan of organization is being used extensively. A unit represents a way of organizing materials and activities for instructional purposes. The basic purpose of unit planning is to relate learnings to one another and to avoid a programme that is highly fragmented by discrete daily lessons.

Meaning of the Term : The term *unit* signifies the unity or wholeness of learning activities relating to some problem or project. This concept seems to be a natural outgrowth of the new psychologies of learning of the early 1900's, which emphasized learning by whole and the importance of the learner as a purposive organism concerned with problems of adjustment in life situations. This development gave importance to integrated learning products as opposed to the atomistic accumulation of facts, abilities and skills. A unit represents not merely the subject matter but also learning experiences. Emphasis is therefore placed upon the organization of courses into units of experience that cut across the subject matter.

A unit may be defined as a large subdivision of a subject with a principle or a topic for its core in which the activities of the pupils are thoroughly planned to give complete mastery of the essentials.

Unit plan should not be confused with the daily lesson plan commonly used to designate daily planning. Daily lesson plans should, of course, be made within the context of the over-all unit plan and should be continuous from one day to the next. That is, the work of each day is an extension and continuation of the previous day's work. A unit may extend over a period of about 8-10 periods, depending upon the subject matter and also the objectives to be achieved.

In planning a unit we have to consider the following aspects :

(a) 'Why' aspect of the unit : You have to state the objectives and to specify them.

(b) 'What' aspect of the unit : You have to do the content-analysis of the unit.

(c) 'How' aspect of the unit : You have to discuss the method and the learning experiences that you will organize.

(d) 'Evidence' for achievement : You have to describe the tool that should be used for evaluation.

If the four aspects mentioned above are attended to, the concept of unit planning will help improve the teaching-learning process. The idea of unit planning was developed because teachers felt that the daily-lesson procedure was fragmentary and inefficient. In other words, the unit concept was a revolt against the daily-lesson procedure. This new concept promotes long range planning and learning in larger wholes that are more meaningful. This concept lays emphasis on the use of various learning experiences for achieving different learning outcomes in the form of understandings, attitudes and skills. One can not assess these outcomes through the fragmentary procedure of daily lessons. The unit plan does not merely rely on the text-book since the subject matter from the text is considered only as a means to an end and not the end in itself.

Units in Specific Subjects

The syllabus of a particular subject can thus be divided into suitable units. The planning of various units will then lead to the year's plan of work. A few examples are given below to show the type of units in different subjects.

1. *Biological Sciences* : Units in this subject should be centred upon some major life process or activity or upon some important aspect of man's relation to his biological environment, for example :
   - (i) Protecting ourselves from disease.
   - (ii) Cells — their structure, function and development in plant and animal life.
   - (iii) The life functions of plants and animals.
   - (iv) How does man control the improvement of living things ?
2. *Physical Sciences* : Units in the physical sciences should be centred upon essential physical processes, principles or phenomena in everyday life, or upon important problems of everyday life involving physical processes, for example :
   - (i) How do machines work for us ?
   - (ii) What conditions determine weather and climate ?
   - (iii) How is electricity produced and transmitted ?
3. *Educational Psychology* : Units in this subject should have a bearing on problems pertaining to class-room teaching for example :
   - (i) What are the factors that influence the behaviour of a child ?
   - (ii) What are the various stages of growth and development ?
   - (iii) What makes learning permanent ?
   - (iv) How can individual differences be assessed ?
4. *Social Services* : Units should be centred upon important and comprehensive aspects of culture, life of people, economic agencies, types of government etc.
   - (i) Early history of the country describing the culture, traditions, beliefs and attitudes of its people.
   - (ii) The Industrial Revolution and its effects on the society.
   - (iii) Evolution of various forms of government.
   - (iv) Influence of other cultures.

5. *English* (Second Language, Std. VII) : Units will be centred upon some structures which will introduce various sentence patterns, for example :

   (i) Tenses.
   (ii) Active and passive voices.
   (iii) Anomalous finites.
   (iv) Verbal nouns, participles.
   (v) Sentence patterns : adverb clauses.
   (vi) Indirect speech.

6. *Mathematics* :

   (i) Fractions in measurement.
   (ii) Mastering the use of decimals.
   (iii) Simple equations.
   (iv) Graphs.
   (v) Transactions in Banks.

7. *Commercial Subjects* :

   (i) Fundamental concepts in book-keeping.
   (ii) Writing of business letters.
   (iii) Using the typewriter.
   (iv) Problems in money transactions.

Successive Steps in Unit-Planning

Any unit must first be seen as a whole, that is, its nature and importance must be determined before any attempt is made to write out the successive steps.

1. *Selection of Objectives* : The first step is to determine the general purpose or goal of the unit. This should be further spelt out in terms of learning outcomes. This is the 'why' aspect of the unit.

2. *Content Analysis* : The syllabi as prescribed at present by the Boards or Departments of Education merely list down various topics. The listing does not indicate the scope of various topics as well as the expected learning outcomes. This practice leaves the teachers to interpret the syllabi in their own way. Most of the teachers depend solely on the text-books prescribed in schools and colleges. These syllabi do not also give enough guidance to the paper-setters and examiners at the public examinations, with the result that teachers, paper-setters and

examiners interpret the syllabus in their own way. If the scope is defined, it will help the teachers as well as the paper-setters. When the scope is not defined, text-book writers also interpret the syllabus in different ways. It is observed that most of the text-book writers, teachers and paper-setters visualise the informational aspect of the subject.

In unit planning emphasis is given on content analysis. This analysis helps the teacher to teach with full awareness of the depth she has to go in respect of each unit. While teaching any unit, a teacher comes across certain new terms, facts, concepts, principles, generalizations etc. She should therefore analyse the content of the unit and list the terms, concepts, generalizations etc., to be discussed. Content-analysis should not be a mere compendium of facts to be learnt. The nature and scope of the unit should be defined. Relative value and importance of essential material should be determined in terms of objectives of the unit.

3. *Learning Experiences or the 'how' aspect* : It is important to select student activities that will best achieve the objectives. While selecting these activities, the question of individual differences must be given careful consideration. Individual differences are reflected not only in the abilities of students but also in their interests, skills and personality traits. In unit planning a teacher will have to provide for these differences.

4. *'Evidence' for Achievement* : In planning any unit, the teacher must give some thought to the type of evaluation tools that she would use for getting evidence. The change in behaviour may be in respect of understandings, skills and attitudes. An attempt to assess the change in these directions must be made by the teacher.

We shall now consider a few units.

[1] English

Unit : Verbal Nouns Standard VII, Periods : 8
(infinitives and participles)

1. *Objectives* : To develop ability
   (i) to identify verbal nouns.
   (ii) to distinguish between common nouns and verbal nouns.

(iii) to use verbal nouns in sentences.
(iv) to locate errors in the use of verbal nouns.
(v) to acquaint with various forms of verbal nouns.
(vi) to classify verbal nouns according to their function.

2. *Content Analysis* : Mention all items that you want to cover under this unit.

(i) *Concept of infinitives* : He wants to *read* this book.
(ii) *Concept of participles* : *Walking* is a good exercise.
(iii) *Various uses of infinitives* :
  (a) As object after verbs : He tries *to speak* English.
  (b) To be glad to .... : I shall be glad *to join* the team.
  (c) As predicative adjunct complement to object : He wants him *to buy* a book.
  (d) Infinitive of purpose : You come to school *to study*.
  (e) Adjectival use of infinitive : There is a house *to let*.
  (f) Infinitive with anticipatory *It* : It is difficult *to speak* Sanskrit.
(iv) *Uses of Participles* :
  (a) Participle phrases : Have you ever seen Rajesh *studying* in the class ?
  (b) Used as subject and object : *Climbing* a tree is a good exercise. Judy likes *swimming*.
  (c) Object of preposition : Barbara is fond of *hiking*.
  (d) Special uses : Mother is going *shopping*.

3. *Learning Experiences* :

1. Study sentences to derive the concept of infinitives.
2. Underline infinitives from given sentences.
3. Compare sentences having nouns and those having infinitives.
4. Classify infinitives according to their function; give various illustrations.
5. Ask pupils to use infinitives in their own sentences.
6. Ask pupils to use infinitives according to specific function.
7. Identify infinitives and participles.
8. Ask the students to correct the errors in given sentences.

This list of activities is not complete in itself. A teacher will have to devise various exercises according to the abilities and needs of pupils.

4. *'Evidence' for Achievement* : This can be obtained through various means such as daily conversation in the class-room, daily assignments and periodical examinations. At the end of the unit, it is desirable to prepare a unit-test. Technique of a unit test will be discussed in the next chapter. After the test is administered, the results should be analysed. The teacher can locate the weaker areas. She can also think of modifying her methods if the results are not satisfactory. She can think of additional learning experiences which will help achieve the objectives.

The unit plan thus shows the inter-relatedness of objectives, learning experiences and the tools of evaluation described in Fig. 1.1 on page 13.

[2] Science

Unit : Thermometers Standard IX, Periods : 4

1. *Objectives* :
   - (a) *Knowledge and Understandings* :
     - (i) To acquaint the pupils with principles underlying measurement of temperature.
     - (ii) To acquaint them with various types of thermometers.
     - (iii) To understand various terms and concepts such as temperature, degree, freezing and boiling points etc.
   - (b) *Skills* :
     - (i) To enable pupils to read thermometers.
     - (ii) To enable pupils to perform experiments involving the use of thermometers.
   - (c) *Applications* :
     - (i) To record daily atmospheric, temperature and to interpret the same.
     - (ii) To interpret readings from a clinical thermometer.

2. *Content Analysis* : What is the scope of this unit ? What are the facts, terms and concepts involved in this unit ?
   - (a) *Facts* : (i) A thermometer registers temperature of a material body. (ii) Mercury is usually used in thermometers. (iii) Specific properties of mercury make it more useful for this purpose. (iv) There are various scales. (v) Clinical thermometer is used to record temperature of human body. (vi) There are many types of thermometers.
   - (b) *New terms* : Temperature, Degree, Fixed points, Scales etc.

(c) *Concepts* : (i) Lower fixed point and upper fixed point. (ii) Degree of hotness of a body. (iii) Heat and Temperature.

(d) *Principles* : (i) Heat expands and cold contracts a body. (ii) Heat flows from higher level to the lower level.

(f) *Symbols and Formulae* :

$$\frac{C}{5} = \frac{F-32}{9}$$

3. *Learning Experiences* : A teacher may demonstrate some experiments or she may allow pupils to perform experiments themselves. Both these types of activities are to be included under this category.

1. Demonstrating some thermometers.
2. Observing the rise in the mercury column when placed in hot water.
3. Drawing a diagram of a thermometer.
4. Observing change in temperature from morning till evening.
5. Comparing readings in Fahrenheit and Centigrade thermometers.
6. Using a maximum and minimum thermometer.
7. Recording the freezing point and the boiling point of water.
8. Maintaining record of daily temperature.
9. Demonstrating that heat flows from higher level to a lower level.
10. Converting readings from one scale to another.
11. Using a clinical thermometer and observing its peculiarities.

A teacher will have to use a number of teaching aids for organizing student activities. Pupils must be given an opportunity to observe, to record and to infer from the observations. Teacher should not dictate any rule or principle. It should be derived by the pupils through experimentation. Collecting data, analysing and then generalizing is a good learning experience for pupils.

4. *Evidence for Achievement* : Has the change in the desired direction taken place? The change in behaviour may be in respect of knowledge, understanding, skill, attitude and interest. For assessing the skill, practical examination in science must be conducted. Let the students read the temperature of hot water and let them record it. Change in attitude and interest can be assessed by observing pupils in various situations. Objectives of knowledges and understandings can be assessed through a unit test. All these methods taken together will give a reliable picture of the pupil's achievement. If the teacher finds that all of her students know the topic but they have not acquired the necessary skill in using thermometers, she will have to modify her approach. She will have to give ample opportunity to pupils for observing and experimenting in the laboratory. Unit plan should not become a rigid one. There should be flexibility in planning. While planning, we may also give introductory remarks which will help the reader to understand the maturity level of the pupil and also the scope of the unit.

[3] Civics

Unit : The Central Government Standard XI, Periods : 12

Sub-unit : The Union Executive.

1. *Introduction of the Unit* : The study of the Indian Constitution in the subject civics occupies an important place. Its study would enable the student to take purposeful interest in the functioning of the administrative machinery and thus help in cultivating the desirable qualities worthy of a good democratic citizen.

The scope of the present sub-unit is confined to giving understanding of the composition, powers and functions of the Central Government. The unit is designed to develop the following major ideas.

1. The Central Government embodies the principles of parliamentary democracy but at the same time contains certain characteristics of a presidential system.

2. The President, though a nominal executive head, possesses in certain cases enormous powers under the constitution.
3. The President has to act according to the advice of the Council of Ministers.
4. Cabinet bears collective responsibility to the Lok Sabha and remains in office till it enjoys the confidence of the Parliament.
5. There is no separation of powers between the legislature and the executive in India.
6. The cabinet is the pivot round which the whole administrative machinery revolves.
7. The Prime Minister is the key-stone of the cabinet arch.

2. *Objectives* :

1. To acquaint the pupils with the salient features of parliamentary and presidential types of Government.
2. To enable pupils to understand the powers of the President such as legislative, executive, judicial, financial and emergency powers.
3. To identify the relationship between the President and the Cabinet.
5. To recognize the relation between the legislature and the executive.
6. To understand the position of the cabinet in the administrative set up in India.
7. To acquaint the pupils with the rights and privileges of the Prime Minister.

3. *Content-Analysis* : What are the various terms, facts and concepts involved in this unit ?

(a) *Facts* : (i) There are three branches of the Government machinery namely the legislature, the executive and the judiciary. (ii) In different countries there are different types of Government. (iii) The prime minister enjoys a privileged position in the country. Ministers of the cabinet are appointed by the President on the advice of the Prime Minister. (iv) When the Prime Minister resigns, all the ministers of the cabinet automatically resign.

(b) *New Terms* : (i) Legislative, executive, judicial, financial and emergency powers. (ii) Democracy. (iii) Parliament. (iv) Council of Ministers.

(c) *Concepts* : (i) Parliamentary democracy. (ii) Presidential type of Government. (iii) Collective responsibility of the cabinet. (iv) Cabinet is responsible to the Parliament. (v) Enjoying the confidence of Parliament. (ci) Separation of legislature and executive.

Students must understand the above concepts which are new to them.

4. *Learning Experiences* : The activities given below are only suggestive and the teacher may select suitable ones as per the needs of her class.

   1. Discussion regarding the nature of Union Government.
   2. Collection of information from newspapers about the elections that might have taken place in the recent past.
   3. Preparation of charts showing the procedure of Presidential election.
   4. A comparative table showing the powers of the President in normal times and in emergency.
   5. Discussion regarding the constitutional position of the head of Great Britain, U.S.A., U.S.S.R. and India.
   6. Discussion regarding the composition, functions and powers of the cabinet.
   7. Organizing a debate on presidential type and parliamentary type of Government.
   8. Organizing mock parliaments to discuss current problems of interest.

9. Formation of a mini-cabinet in schools to inculcate qualities of leadership, spirit of tolerance, and of co-operation, respect for others etc.
10. Listening to the commentaries on parliamentary proceedings broadcast by AIR.
11. Holding elections in the school, allowing some time for election-campaign (Fig. 5.4, page 90).
12. Arranging film-shows on topics relevant to the subject.

5. *Evaluation* : The desired change in attitude and the development of qualities cannot be assessed through a written examination. The teacher will have to use some other tool such as a check-list and observe the pupils' behaviour in day-to-day situation in the class-room or the play-ground or in other social situations. The objective related to knowledge and understanding may be assessed with the help of a unit test.

If the teacher finds that desired qualities are not inculcated among the students, she will have to emphasize this aspect in day-to-day situations.

The qualities such as tolerance, respect for others, initiative can be seen when students participate in such activities as debating, sports-competitions, annual-socials, and such other group activities. Teacher must however be watchful. Use of other tools such as rating scales, check-lists will be discussed in Chapter 12.

So far we have discussed unit planning in a running format. We can conveniently use a tabular form for this purpose. Two more units have been illustrated with the help of a chart on the following pages.

A teacher should prepare the daily lesson-plan, only after completing the *unit-plan.* It should be made within the context of the over-all unit plan. If there are six lessons in a unit, there should be continuity from one lesson to another. The work of each day is an extension and continuation of the previous day's work.

Year's Plan of Work

When the planning of various units in a subject has been completed, we can prepare the year's plan. In preparing the year's plan it is desirable to show the time allotment to various units. It is also desirable to show the weightage in terms of teaching periods that the teacher wants to give to various objectives. Let us consider the year's plan for the teaching of Geography in standard V. Units 2 to 5 are to be discussed not with reference to India but with reference to the particular State such as Mysore or Andhra or Gujarat or Maharashtra (Page 117).

Interpreting the Year's Plan

The plan has been prepared with some assumptions : There are two periods allotted to the teaching of Geography per week in standard V. During the whole year there will be about 32 weeks available for actual teaching work. There would thus be 64 periods available for the teaching of Geography.

While teaching, the teacher should not rigidly follow the distribution of periods objectivewise, for example in unit 7 the weightage to three objectives is in the ratio of 5 : 3 : 2. A teacher should not say, "I shall achieve the first objective during the first five periods, the second objective during the next three periods and the third objective during the last two periods." It would be a mockery of the unit plan. The distribution only shows the weightage the teacher should give to various objectives in teaching the unit. Sometimes the teacher may attend to all the objectives in a single period while on other occasions she may attend to only one or two objectives.

No separate periods have been shown for developing interests and attitudes because it is expected that these are the natural outcomes of learning that may take place during the course of a year. These are as if the by-products of learning. A teacher should not say, "Let us spend some time to-day for developing interest in Geography." Such an approach will defeat the purpose.

Validity of the Year's Plan : As the saying goes, "The proof of the pudding lies in its eating," so also the year's plan is

Table 6.1

Year's Plan

Standard V·
Subject : Geography

Periods per week : 2
Total periods : 64 (approx.)

No. of Periods

Unit	Knowledge and Understanding	Application	Skill	Interest and attitude	Unit Test	Total
1. Revision of previous portion	1	2	1	✓	1	5
2. Agriculture in Maharashtra or Mysore or the concerned State	2	1	1	✓	1	5
3. Occupations and industries in the concerned State	2	1	-	✓	1	4
4. Forest, Mineral and Ocean wealth in the State	2	1	1	✓	1	5
5. Means of Transport in the State	2	1	1	✓	1	5
6. India : physical features, climate. vegetation	3	1	2	✓	1	7
7. India : communication, industries and occupations	5	3	2	✓	1	11
8. Physical Geography : factors determining the climate	3	1	1	✓	1	6
9. The Earth : shape, size, rotation of the Earth, latitudes and longitudes	3	2	1	✓	1	7
10. Practical Work : (i) Visits to fields, factories, market places (ii) Recording weather-changes, rainfall, position of the Sun, use of compass	3	3	2	—	1	9
	26	16	12	—	10	64

to be tested through implementation. The distribution given in the plan is tentative. A teacher should record the actual time required for completing the various units, and the weightage given to various objectives. Then at the end of the term or the year, she can test the validity of her plan. Actual teaching will supply the feed-back material. She can then modify the next year's plan, if necessary.

When any unit has been completed, the teacher should try to get evidence regarding the achievement of objectives. This is done with the help of a unit test. We shall discuss this phase in the next chapter.

FOR FURTHER STUDY

1. What is a Unit Plan? Discuss the various steps in Unit-Planning.
2. Select any one unit from social studies, mathematics, English, science or the mother tongue for standard VIII and give a detailed plan of the same.
3. What is meant by content analysis of a unit? Illustrate by selecting a unit from any subject of standard IX.
4. Prepare a year's plan for teaching any one subject of standard X. How will you test the validity of your plan?

* * *

5. Mathematics

Subject : **Algebra** Standard VIII, Periods : 7.
Unit : Simple equations and problems based on them.

Sub-unit	*What aspect*	*Why aspect*	*How aspect*	*Evidence*
	Content Analysis (teaching points)	Objectives with expected learning outcomes	Learning Experiences	Tools of Evaluation
1. Introducing Simple Equations	*Terms* : equation, unknown, anxioms. transposition.	*Knowledge* : Pupil recalls the terms.	Teacher uses the two pans of a balance and introduces the idea of equation.	(i) Oral test at the end of the lesson. (ii) Solving examples in the classroom. (iii) Daily assignments. (iv) Unit test may be given at the end of the unit.
	Signs and *Symbols* : = equal to > greater than < less than	*Understanding* : Pupil explains the terms and symbols. *Application* : (i) He uses symbols in various situations. He symbolises verbal expressions.	Pupil tries to balance the two pans by using different weights. Teacher fixes the terms through drill work.	
	Principles or Axioms : Equation remains unaltered if the same or equal quantities are added or subtracted from both sides.	(ii) He uses axioms in solving the equation.	Oral practice in testing the axioms.	
	Procedure :	Pupil detects errors. He explains the procedure.	Practice in solving the equation.	

	(i) Solving the equation. (ii) Verifying the solution.	He explains the procedure. He uses proper terms and symbols. He does it with proper speed.	Pupil verifies the solution. Practice for developing skill.	
2. Solving Problems	*Procedure* : (i) Framing the equation. (ii) Solving the equation. (iii) Verifying the solution.	*Understanding* : (i) Pupil verbalises the data and expresses in terms of equality. (ii) solves for x, the unknown. (iii) detects errors. *Application* : (i) Translates a problem into an equation. (ii) Verifies the solution.	(i) Teacher demonstrates the procedure of translating a problem in the form of an equation. (ii) Drill work in framing the equations. (iii) Solves some problems in the class-room. (iv) Explains the method of verifying the results. (v) Drill work in solving equations. (vi) Asks pupils to detect errors in the solution.	(i) Oral work. (ii) Daily assignments.
3. Constructing problems from given equations	*Procedure* : (i) Verbalising an equation. (ii) Converting it in the form of a problem. (iii) Constructing problems.	*Knowledge* : Reading an equation. *Understanding* : Verbalizing an equation e.g. when three are added to x, we get seven. Writing it in the form of a problem. *Application* : Prepares a problem from a given equation.	Teacher demonstrates the method. Pupil does the work of verbalising the equation and converting it in the form of a problem. Detects errors in the wording of the problem.	Unit test.

7

UNIT TEST

We have discussed in the previous chapter various steps in unit planning. The last stage in unit planning is the preparation of a Unit Test. It is a small test to be given at the end of teaching — a given unit. A unit test is not just a random assortment of questions. To be an effective instrument of evaluating academic achievement, it has to be structured according to a pattern decided in advance. Following steps are necessary for setting up a good unit test.

A. Preparation of a Design

The design is not only the first but is also the most important step in setting any question paper. It lays down the chief dimensions of the test. These dimensions are as follows :

1. *Selection of Objectives* : There are certain objectives which the teacher expects to achieve through her teaching. The test should aim at assessing the achievement of these objectives specifically. Suppose a teacher sets out to achieve the following objectives in the teaching of science in standard VIII :

(i) Knowledge of facts, concepts, principles etc.
(ii) Application of scientific knowledge to life situations.
(iii) Experimental and practical skills.
(iv) Scientific interests.

The teacher might have emphasized objectives (i), (ii) and (iii) in her teaching. Then she should plan to test only these three at the end of the unit. She may consider that objective (iii) will be assessed by observations of the pupils in the laboratory, hence she may select only the first two objectives for her unit test. Specifications of the objectives are given in the unit plan which should be considered while framing the questions.

2. *Weightage of Objectives* : This step is important in order to decide the number of questions to be provided to meet any one objective in relation to others. In other words the relative importance of each objective must be considered. If the teacher thinks that the *knowledge* objective has been emphasized more than the *application* objective, she may choose to have 60% of the questions on the first and 40% on the second objective. The weightage given to the two objectives in this case is thus in the ratio of 3 : 2. It should be realized that the overall emphasis given to different objectives should be properly decided by a committee of experts and should not be left completely to chance or to individual discretion of teachers.

3. *Weightage to different areas of content* : This entails the analysis of the unit and the delimitation of the scope of each sub-unit. In assigning weights to various areas of the content, a common procedure is to weigh each area in terms of the time devoted to it. This weightage should be reflected in the number of questions included in the test. This aspect is important when the question paper is to be based on a large content area consisting of 5-6 units.

4. *Weightage to different forms of questions* : Working on the principle that for testing a particular ability and content, the most suitable form of question should alone be used, we may like to choose different forms of questions instead of having just the traditional form namely, the essay type question. Having made this decision, the marks to be allotted to each form will have also to be decided. The essay type, the short answer type and the objective type of questions are useful in testing various learning outcomes. If the weightage to these forms is in the ratio of 3 : 3 : 4, marks allotted to these forms would be 30, 30 and 40 respectively.

5. *Scheme of Options* : The design of the test may also indicate the pattern of option, that is, complete elimination of overall options and retention of internal options. The optional questions should be comparable in respect of the objectives to be tested, the form of the question, the major area of content and the difficulty level of the question. When overall options are given.

pupils answer different questions and the common basis for evaluating their achievement is lost.

6. *Sections in the Question Paper* : The inclusion of objective type questions may necessitate the division of the question paper into sections and the allotment of a separate time limit for answering each. These questions are usually answered in a shorter time and need not remain with the pupils for longer than just the time necessary for attempting them. In view of this, the design may indicate the division of the question paper into two sections, the one containing the objective type questions may be collected at the end of the given time limit and the second may be administered after a little break.

All these points have been considered while discussing the design of a test at a later stage in this chapter.

B. Preparation of a Blue-Print

The blue-print gives the details of the design in concrete terms. It is a three-dimensional chart giving the placement of the different questions in respect of :

(i) the objectives tested by each.
(ii) the content area covered by each.
(iii) the form of questions which is most suitable for testing (i) and (ii) above.

If a chart specifying only the first two aspects mentioned above is prepared, it will be a two-dimensional chart. The blue-print is thus a detailed plan based on the design for preparing a question paper. It serves the paper-setter in much the same way as a blue-print serves the carpenter. It helps him in obtaining the finished product as per specifications. In the case of the test maker, it provides greater assurance that his test will measure the learning outcomes and the course content in a balanced manner. The blue-print of a question paper has been discussed at a later stage in this chapter.

C. Preparation of Questions based on the Blue-Print

After the position of different questions is located on the blue-print, the next step is the preparation of questions, the dimensions of each of which have already been defined on an individual basis in the blue-print. Every question is to be framed with reference to the objective and the content area indicated in the blue-print. While writing down a question, following criteria of a good question should be attended to.

(i) It should be on a specific objective.
(ii) It should relate to a specific content area.
(iii) It should be written in the form as required by the blue-print and should satisfy the rules for framing that particular form of question.
(iv) It should promote thinking rather than rote memorization.
(v) It should be at the desired level of difficulty.
(vi) It should clearly indicate the scope and length of the answer.
(vii) It should sharply distinguish between the brighter and dull pupils.

The last criterion will not apply to speed tests or to diagnostic tests. Maxims for writing the various types of questions will be discussed in chapter 11.

D. Editing the Question Paper

Following measures should be taken while editing a test.

1. *Grouping the Questions* : This is done usually on the basis of the form of questions. Essay type and short answer type may be grouped in one section while the objective type questions may be grouped into another section. Questions in each section should be arranged in order of difficulty from the easy to the more and more difficult. In the section of objective type questions, items of the same form should be grouped together because each form requires separate instructions to the students.

2. *Instructions to Examinees* : Directions regarding what the students are required to do, need to be clear, specific and pointed. General instructions may be given at the beginning of the paper and specific instructions related to each section may be given at the beginning of the section. Marks allotted to each question or its parts should be indicated on the question paper. If students are expected to draw a diagram, it should be clearly indicated.

Instructions regarding options, length of expected answers and time allotted should be indicated.

3. *Implications for Administration* : Editing may have important implications for administration of the question paper, for example, it may be desirable to administer the two sections separately with independent time limits to each.

E. Preparation of the Marking Scheme and the Scoring Key

Marking scheme is particularly essential in the case of essay type and short answer type questions. It gives the expected outline of the answer and the marks that each point or aspect of the answer deserves. We may indicate the number of points expected in the answer and the weightage to these points. Marking scheme is very essential when many examiners are involved in assessing the answer-books. All of them must know the paper-setter's view point regarding expected answers and allotment of marks. The marking scheme helps the examiners to bring about uniformity of standards in marking. In objective type questions, this problem does not arise since answers are definite and generally equal marks are allotted to all items in the test.

Scoring key should be prepared in case of objective type questions and a marking scheme for the essay type and short answer type. Very often we discover mistakes in the content or the wording of the questions when we start preparing the scoring key and the marking scheme. Hence it is desirable that these two are prepared simultaneously with the question paper.

F. Reviewing the Question Paper

It is a good practice not to hand over a question paper immediately as it is set, but to review it once again before it is sent to the press. This reviewing should be done after a little interval. On a fresh morning the teacher may go through the paper once again. Quite a few things may strike him on this second reading. An improvement in the wording of a question may be possible or a slight alteration might remove some ambiguity which is lurking at some point in a question. A teacher can thus avoid the unhappy experience of discovering certain flaws in the paper in the examination hall.

G. Questionwise Analysis

This analysis helps us to know the strengths and weaknesses of the question paper more thoroughly. It will also help us to tally the question paper with the blue-print. Each question paper will be analysed in terms of the following elements :

(i) Objectives tested by various questions.
(ii) Specifications on which the question is based.
(iii) Topics covered by the questions.
(iv) Form of the question.
(v) Estimated difficulty level.
(vi) Approximate time required for answering.
(vii) Marks allotted to each question.

These details are given in a tabular form and each column of the table can be summarized to give us the distribution of question either over the different objectives or the areas of content or the forms of questions etc. This analysis is given in the last section of this chapter (page 133).

This analysis helps us to determine the *content validity* of the test. This is done merely by inspection of the various questions or items. But this analysis cannot be considered as valid unless it is proved experimentally. We have therefore to consider the test-result and to analyse the answers given by the pupils.

H. Item-Analysis

A teacher ordinarily thinks that her job is finished when she has prepared, administered and scored the test. If she stops at that she loses much of the value of the test.

How effective is the test ? : The taste of the pudding lies in its eating. Similarly effectiveness of the test can be judged only after using it. Responses of the examinees are to be analysed. Items which are effective are retained while those which are not effective either to be discarded or to be improved. This is known as the process of item-analysis. This step is usually followed in the process of standardization of a test. We shall discuss this aspect in the next chapter.

Various steps mentioned above should be considered not only while preparing a unit test but also while setting the question paper of the annual examination or the Board examination. If the design of the Board examination is known to the teachers, it will help improve the teaching process.

We shall now consider a sample question paper along with its design and blue-print. Questions based on the language-text have not been included in this paper since a reference to various lessons of the text may not be meaningful to all readers.

A Sample Question Paper in English for Standard VIII.

Design of the Question Paper.

Table 7.1 : Weightage to Objectives

Objective	Marks	Percentage
1. Knowledge	22	22
2. Comprehension	28	28
3. Expression	50	50

Table 7.2 : Weightage to Content Area

Area	Marks	Percentage
1. Elements of language	38	38
2. Unseen passage	12	12
3. Composition	50	50

Table 7.3 : Weightage to different Forms of Questions

Form of Question	No. of questions	Marks
1. Essay	3	36
2. Short answer	5	24
3. Objective type	15	40

Blue-Print of the Question Paper

Table 7.4: A three dimensional chart

Objectives	*Knowledge*			*Comprehension*			*Expression*			*Total Marks*
Form of Questions Content area	**E**	**S**	**O**	**E**	**S**	**O**	**E**	**S**	**O**	
1. Elements of language	—	10	28	—	—	—	—	—	—	38
2. Unseen passage	—	—	—	—	—	12	—	—	—	12
3. Composition	—	—	—	—	—	—	36	14	—	50
Total marks	—	10	28	—	—	12	36	14	—	100

E = essay type, **S** = short answer, **O** = objective type.

The blue-print is a three dimensional chart; it describes the three dimensions of the question paper namely the objectives, the content area and the form of questions used. In Table 7.4, we find that the three objectives have been given the weightage in the ratio of 38 : 12 : 50. Marks are allotted to the three content areas in the ratio of 38 : 12 : 50 while the three forms of questions E, S and O are in the ratio 36 : 24 : 40.

There are two sections and all questions are compulsory.

A Sample Question Paper in English

Time : 2 hours, *Standard* VIII *Marks* : 100

General Instructions

1. This paper has two sections **A** and **B**.
2. Answers to section **A** should be written on the question paper itself. Answers to section **B** should be written in a separate answer-book.
3. Attempt section **A** first. Time for section **A** is 30 minutes. In case you finish it earlier, you may pass on to section **B**.
4. All questions in section **A** and **B** are compulsory.

Section A

Time : 30 minutes. Marks : 40

Instructions for Section A :

(i) This section will be collected at the end of 30 minutes.

(ii) If you find any question difficult, do not waste your time on it but proceed to the next question and come back after other questions are answered.

(iii) Each item in a question carries one mark.

Q. 1. Fill in the blanks with a suitable word given in brackets :

(i) The typewriter is .... than this pencil.
(light, heavy, lighter, heavier)
(ii) This wall is just as .... as that one.
(thick, thickness, width, height)

Q. 2. Use the proper form of the verb :

(i) I do not know how many pounds I ....
(weighing, weigh, weighs)
(ii) Mike .... heavier than his sister Barbara.
(am, is, are)

Q. 3. Join the two sentences by choosing a suitable word given in the parenthesis :

[but, where, because, when]

(i) Susan went to the seaside. She bathed.
(ii) Harish slept until seven. The alarm clock went off at six.
(iii) Judy was waiting at the station. The bus arrived.

Q. 4. Fill in the blanks with suitable words given in the bracket:

[wait for, turn on, pick up, taking back]

(i) Would you mind .... these books .... to the library ?
(ii) Please don't .... me. I will be busy all afternoon.
(iii) Please ask Linda to .... the lights.

Q. 5. Underline the word that is different from the remaining four words in each line.

(i) summer, winter, easter, spring, autumn.
(ii) run, skip, hop, jump, stand.
(iii) inch, gramme, yard, mile, foot.

Q. 6. Fill in the blanks with a suitable word given in the bracket : [aunt, uncle, brother, nephew, husband]

(i) My brother's son is my ....
(ii) Yasmin's uncle is her father's ....
(iii) Gloria's father is her mother's ....

Q. 7. Use a suitable preposition in the blank space :

[on, from, of, except]

(i) All have passed the examination .... Raju.
(ii) There has been a lot .... improvement here.
(iii) Mukesh spent his childhood .... a farm.

Q. 8. Select a suitable abstract noun from the given words :

[size, height, weight, length, shape]

(i) The .... of the table is round.
(ii) The .... of the bag is 10 kilogrammes.
(iii) The .... of the street is two miles.

Q. 9. Use one word for the group of words in italics. :

(i) Meena came in *without making a noise.* ( )
(ii) Kusum pays for her milk *every day.* ( )

Q. 10. Underline the word from the parenthesis, that has nearly the same meaning as the word outside the parenthesis :

(i) large .... [top, long, great, wide]
(ii) joy .... [help, happiness, sleep, noise]

Q. 11. Underline the word from the parenthesis that is opposite in meaning to the word given outside the parenthesis :

(i) top .... [middle, down, bottom, below.]
(ii) brave .... [weak, timid, strong, quiet.]

Read the following story and answer the questions that follow :

Once upon a time, while a wolf was hastily eating a hearty meal, he got a bone stuck across his throat. He tried and tried but he could get it out. It would *neither* move up nor *down.* He was in great pain and ran about howling and offering a handsome reward to any one who would help him.

Soon he met a crane who felt very sorry for him. She promised to *pull* the bone from his throat. To do this, she stretched out her *long* neck and reached into the wolf's mouth with her long bill.

In a *few* minutes the painful bone was gone and the wolf leaped to his feet and started to walk away. "But what about my reward!" cried the crane.

The wolf whirled about angrily and growled, "Reward! Why, you greedy creature, what reward do you expect? You have already had enough *reward*, for you had your head in a wolf's mouth without having it snapped off."

And so the crane learnt that you may expect no reward for helping the wicked.

Answer the following questions :

Q. 12. Draw a line under the right ending :

(1) The wolf asked for
(i) a doctor (ii) help (iii) drink.

(2) The wolf was eating
(i) slowly (ii) hastily (iii) nicely.

(3) The crane's neck was
(i) long (ii) thick (iii) short.

Q. 13. Write after each line the answer *Yes* or *No*.

(1) The crane reached into the wolf's mouth. ....

(2) The crane said the wolf was greedy. ....

(3) When the bone was gone, the wolf was still in pain. ....

Q. 14. Draw a line around the correct ending :

(1) The crane in the story was
(i) a lifting machine (ii) a bird
(iii) a coloured pencil.

(2) A hearty meal would be
(i) a big meal (ii) a meat meal (iii) a small meal.

(3) To put her head into the wolf's mouth was
(i) wise (ii) natural (iii) dangerous.

Q. 15. Choose the words in italics from the story which are the opposites of:

(i) short ...... (ii) many ...... (iii) push ....

Section B

Time: 90 mintes Marks: 60

Instructions :

(i) Answer the questions in a separate answer-book.
(ii) Answers should be precise and to the point.
(iii) Figures to the right indicate full marks.

Q. 16. Punctuate the following sentences: (4)

(i) the train stopped the passengers got out quickly
(ii) when will you return to bombay asked mohan

Q. 17. Change into indirect: (2)

Peter said to his father, "I want to purchase this book."

Q. 18. Use the following phrases in sentences of your own. (8)

(i) in front of (ii) to pay attention
(iii) in spite of (iv) in time

Q. 19. Bring out the difference in the following pairs of words by using them in sentences: (6)

(i) tale, tail (ii) write, right (iii) advice, advise.

Q. 20. Develop the following points into a story: (12)

A fat man .... visits a friend .... friend's dog chases him .... he runs .... friend shouts .... barking dogs never bite.

Q. 21. Rewrite the following paragraph changing 'I am' to 'Gopal is'. Make any other changes which are necessary: (4)

I am a student of this school. I have been here since July. I didn't like it at first, but now I am getting used to it. I have to attend the school for five hours. Half of my time is spent in the classroom. For the rest of the time, I play games or read in the library.

Q. 22. Write in about 15 lines a story as seen in the following pictures. Write it in the past tense. (12)

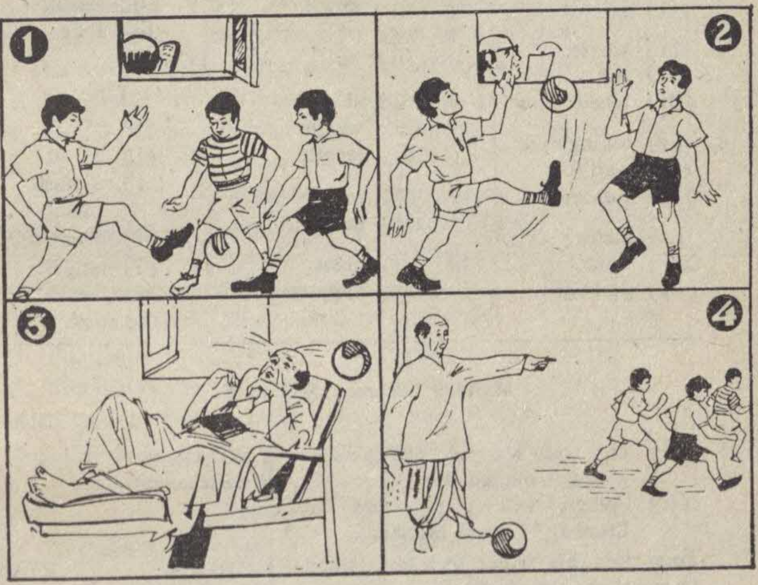


Fig. 7.1 : Picture Composition.

Q. 23. Suppose you are staying in the hostel of your school. Write a letter to your father requesting him to allow you to join the school trip to Kashmir. (12)

[**N.B.** It should be remembered that it is not necessary to include all the above types of questions in one unit test. The number of questions will obviously depend upon the time available for the unit test. Here an attempt is made to show the variety of questions that can be used for assessing various objectives. Questions based on the text-book should also find a place in the test.]

Scoring Key : Section A

One mark is allotted to each item under each question

Q.	*Answer*
1.	(i) heavier (ii) thick
2.	(i) weigh (ii) is
3.	(i) where (ii) but (iii) when
4.	(i) taking back (ii) wait for (iii) turn on
5.	(i) easter (ii) stand (iii) gramme
6.	(i) nephew (ii) brother (iii) husband
7.	(i) except (ii) of (iii) on
8.	(i) shape (ii) weight (iii) length
9.	(i) quietly (ii) daily
10.	(i) great (ii) happiness
11.	(i) bottom (ii) timid
12	(i) help (ii) hastily (iii) long
13.	(i) yes (ii) no (iii) no
14.	(i) a bird (ii) a meat meal (iii) dangerous
15.	(i) long (ii) few (iii) pull

Marking Scheme : Section B.

			Total
16.	(i) The train stopped. The passengers got out quickly. (ii) 'When will you return to Bombay ?' asked Mohan.	½ mark allotted to each point	4
17.	Peter told his father that he wanted to purchase that book.	½ mark for each change	2
18.	Full credit should be given for correct use only.	2 marks for each	8
19.	Full credit should be given for correct use only.	2 marks for each pair	6
20.	Following aspects should be considered: title (1), organization (4), vocabulary (2), structure (2), spelling (2), punctuation (1).	marks as shown in the parenthesis	12
21.	There are 12 changes to be carried out.	1 mark for 3 correct changes	4
22.	Following aspects should be considered: title (1), organization (4), vocabulary (2), structure (2), spelling (2), punctuation (1).	marks as shown in the parenthesis.	12

23.	Following aspects should be considered: address (1), salutation (1), organization (3), spelling (2), punctuation (2), relevance (3).	marks as shown in the parenthesis.	12

Questionwise Analysis

Sr.No.	Objective	Specification	Topic	Form	Difficulty Level	Marks
1.	Knowledge	identifies	vocabulary	O	C	2
2.	Knowledge	recognizes	structure	O	C	2
3.	Comprehension	transforms	transformation	O	B	3
4.	Comprehension	discriminates	vocabulary	O	B	3
5.	Comprehension	discriminates	vocabulary	O	B	3
6.	Comprehension	interprets	vocabulary	O	A	3
7.	Knowledge	recognizes	prepositions	O	C	3
8.	Knowledge	recognizes	abstract nouns	O	C	3
9.	Knowledge	identifies	vocabulary	O	B	2
10.	Comprehension	discriminates	vocabulary	O	B	2
11.	Comprehension	discriminates	vocabulary	O	B	2
12.	Comprehension	locates	unseen passage	O	B	3
13.	Comprehension	interprets	unseen passage	O	B	3
14.	Comprehension	identifies	unseen passage	O	B	3
15.	Comprehension	locates	unseen passage	O	C	3
16.	Knowledge	recalls	punctuation	S	B	4
17.	Knowledge	recalls	narration	S	B	2
18.	Expression	uses	sentence formation	S	B	8
19.	Expression	gives meaning	vocabulary	S	B	6
20.	Expression	organizes ideas	composition	E	A	12
21.	Knowledge	recalls	structures	S	C	4
22.	Expression	organizes ideas	composition	E	A	12
23.	Expression	writes a letter with correct structures	composition	E	A	12

Estimated Difficulty Level: A = difficult, B = average, C = easy.

Form of Questions: O = objective, S = short answer, E = essay.

FOR FURTHER STUDY

1. Discuss the importance of planning a unit test.
2. Explain the term, 'Design of a test.' Discuss the various dimensions to be considered while preparing a design of a test.
3. State the criteria of a good question.
4. Select any one subject and review the question papers of the last annual examination of standard IX of your school with reference to the objectives, content area and form of questions. Discuss the findings.
5. Prepare a unit test in any two school subjects of your choice.
6. Prepare a unit test along with design and blue-print on any one of the following units :

   (i) Simple equations (ii) Transference of heat (iii) Photosynthesis
   (iv) Revolution of the Earth — seasons and their causes.
   (v) Noun clause, adjective clause and adverb clause.
7. Illustrate how the marking scheme helps in maintaining uniformity of standards in marking.
8. Give a question-wise analysis of the test you have prepared under question (6) above.

* * *

8

STANDARDIZATION OF A TEST

We have seen in the previous chapter that the class-room examination needs planning. A mere collection of questions, whatever their number and individual quality, does not make a good test. A unit test will become an effective measuring instrument if the test is prepared according to the blue-print. Such a test will help the teacher in evaluating the progress of her pupils. But it may not help her for comparing the achievement of her pupils with that of pupils from other schools. In such a situation we would prefer a standardized test.

What is a standardized test? : A standardized test differs from an ordinary test in that it is constructed according to some well defined procedure, is administered according to definite instructions, scored according to a definite plan and that it provides a statement of norms. A test is standardized in respect of content, method of administration, method of scoring and setting up of norms. Standardization is a process for refining a measuring instrument through scientific procedures. Many of the steps followed in planning a unit test and described in the previous chapter are to be followed in the process of standardization. We may summarize the various steps here.

[1] Design of the Test

(a) *Defining the purpose of the test* : A test may be used for diagnosing pupil's weak points, or it may be used to assess the achievement of students. Sometimes it is used as a screening device for selecting the best candidates for admission to a medical or an engineering course. Sometimes it is also used as a speed test to know how far the fundamental skills have been mastered by the students. The content of all the above types of tests will differ from one another. It will depend upon the purpose.

(b) *Defining the Objectives* : When the purpose has been defined, it should further be spelt out in terms of specific behaviour changes expected in the pupils. The function to be measured should be analysed. If you want to prepare a standardized test for reading-comprehension, this function should be spelt out in terms of learning outcomes.

(i) What is it that a child with this ability does ?
(ii) How does such a child differ from another who has not developed this ability ?

Answers to these questions will help to get a list of learning outcomes. Questions will have to be framed to cover these learning outcomes. For a discussion of these outcomes readers may refer to page 28.

If you intend to prepare an intelligence test, first define and analyse this function namely *intelligence* into its components.

In what respects does an intelligent child differ from a dull child ? The function can be defined in terms of expected pupil behaviour for example, an intelligent child grasps concepts quickly, understands more complex ideas, perceives finer differentiations in situations, shows originality, uses language more appropriately and precisely and learns more rapidly.

(c) *Determine the weightage to be given* : If there are more than one objective to be tested, the weightage to these objectives will have to be decided. Weightage to various content areas also needs to be determined.

(d) *Preparation of a Blue-Print* : This is to ensure proper coverage of objectives on the one hand and the subject matter on the other hand. We thus need a two-dimensional chart. In standardized tests we do not use essay type or short answer type of questions but only the objective type is used. This third dimension namely the form of questions is not considered. Hence we need only a two-dimensional chart and not a three dimensional chart which is used in a unit test.

[2] Preparation of the Preliminary Form of the Test

(a) *Writing of Items* : Items should be selected in such a way that the expected pupil behaviour in different situations is reflected in them. In a reading comprehension test for example, items should be so selected as will test pupil's ability to grasp the meaning of words, to locate significant facts and ideas from a given passage, to locate key words etc. In an intelligence test items testing the pupil's ability to solve problems through verbal and numerical reasoning will have to be prepared. Each item should fulfil the criteria of a good question discussed on page 122.

(b) *Arrangement of Test Items* : Following considerations must be made while organizing and arranging the items of a test. (i) Items of the same form such as multiple choice, or alternate response type should be grouped together as each set of items requires separate instructions for the students. (ii) Items should be arranged in order of difficulty from the easy to the more and more difficult. (iii) Too many different forms of items may not be used in the same test as it involves some waste of time over reading instructions etc., each time a fresh form is encountered.

(c) *Item Difficulty* : Adjusting the level and range of difficulty of test-items requires proper judgment on the part of the test constructor. The difficulty of the items should be determined by the purpose of the test. If the purpose of the test is to judge the speed or to judge the mastery of the fundamentals of any subject, it will not matter if the items are easy and all of the pupils receive high scores. If however, the test is given to rank the students according to their achievement, it will be necessary to include items of varying difficulties. Such a test should provide suitable opportunity of effort for both the bright and the weak students in the group. For motivational purposes the first few items should be so easy that no body would miss them. The items should cover a wide range of difficulty so that each pupil is allowed to work long enough to reach the level at which his abilities are taxed to the utmost. These are however, matters, in which a teacher gains competence through experience. A numerical index showing the difficulty of items is found out after the preliminary try out of the test is given to a representative sample.

(d) *Preliminary Try Out* : The test thus prepared may be administered to a small but representative group of pupils. This is necessary for collecting preliminary statistical evidence about the individual test items and about the test as a whole. For purposes of editing and revision of various test-items, data are required on the difficulty level and the discriminating power of the test items. For item analysis it is essential that the try out is given with ample time limit so that most individuals will have a chance to try all items. Since this stage is to serve as the basis for item selection, more items must be included in this preliminary form of the test than are to be included in the final form of the test. The items that do not fulfil the criteria can then be either discarded or improved. The merit of a test ultimately depends upon the individual items which make up the test. It is therefore necessary to analyse the items after the preliminary try-out. Time taken by various students to complete the test should be recorded. This will help to fix up the time limit for the final form of the test.

[3] Item Analysis

A teacher ordinarily thinks that her job is finished when she has prepared, administered and scored the tests. If she stops at that, she loses much of the value of the test. By discarding the test in this way, she has little evidence concerning the quality of the test, she used to measure the achievement. A test, apart from serving the purpose of appraising the progress of students, has a diagnostic function to perform. The analysis of the responses helps the teacher to study the learnings of the class and its failure to learn certain things. On the basis of that, she can improve teaching and also guide students to improve their learning. The analysis of responses has also another advantage that of improving the test items themselves and thereby preparing better tests. The effectiveness of each item can be determined by analysing the pupils' responses to the item. This item analysis is usually designed so that answers to the following questions can be obtained.

(1) How difficult is the item ? That is, what is its difficulty value ?

(2) How well does the item discriminate between high and low achievers ? That is, does the item have any discriminative value ?

(3) How effective is each of the distractors in the item ?

(a) *Procedure for Item Analysis* : For most class-room tests a simplified form of item analysis is all that is necessary or warranted. A suitable procedure is to compare the responses of pupils ranking in the upper and lower quarters of the class on the basis of the total score. The responses of the pupils in the middle 50 percent of the group are not included in the analysis but are assumed to follow the same trend as those in the upper and the lower quarters. To illustrate the method of item analysis, let us consider a class of 80 pupils, answering a test in Geography; our item analysis might then proceed as follows :

(i) Rank the papers in order from the highest to the lowest score.

(ii) Select the 20 papers with the highest scores and the 20 papers with the lowest scores, that is, count off 25 percent of the best and 25 percent of the poorest papers.

(iii) For each test item, tabulate the number of pupils in the upper and lower groups who selected each alternative. This tabulation can be made on the test item card as shown in Fig. 8.1.

(iv) Estimate the difficulty of each item.

(v) Estimate the discriminating power of each item, (i.e. difference between the upper and lower groups who got the item right).

(vi) Evaluate the effectiveness of the distractors in each item (i.e. attractiveness of the incorrect alternatives).

Course : Geography Date :
Content : Weather Instruments
Objective : Knowledge

Item : Which of the following instruments is most useful in weather forecasting ?
(A) Anemometer (B) Barometer (C) Thermometer (D) Rain Gauge.

Item Analysis Data

Alternatives	A	(B)	C	D	Omit	Total
Upper 25%	0	15	3	2	0	20
Lower 25%	4	7	4	5	0	20

Difficulty Value = 55% Discriminating Power = 0·40

Fig. 8.1 : A Test item card with item analysis data.

Comments : 15 pupils in the upper group and 7 pupils in the lower group selected the correct alternative B. This makes a total of 22 pupils who got the item right, indicating that the item has a moderate level of difficulty. Since more pupils in the upper group than in the lower group got the item right, it is discriminating positively. That is, it is distinguishing between the high and the low achievers, (as determined by the total test score). Since all the alternatives were selected by some of the pupils in the lower group, the distractors appear to be operating effectively.

(b) *Estimating Item Difficulty* : The difficulty of a test item is indicated by the percentage of pupils who get the item right.

$$\text{Difficulty} = \frac{R}{N} \times 100 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where R = number of pupils who got the item right.
N = total number of pupils in the two groups U and L.

Applying this formula to item analysis data in Fig. 8.1.

$$\text{Difficulty} = \frac{22}{40} \times 100 = 55 \text{ per cent.}$$

An item passed by 80% of the upper group and by 70% of the lower group has a difficulty index of 0·75 that is, $\left(\frac{\cdot 80 + \cdot 70}{2}\right)$ and is considered as an easy item while an item passed by 15 percent and 5 percent of the two groups has a difficulty index of 0·10 which is a difficult item. This estimate of difficulty though obtained from the two extreme groups is sufficiently accurate for class-room use and is easily obtained from the test item card.

(c) *Discriminating Power of an Item* : The discriminating power of a test item refers to the degree to which it discriminates between the bright and the dull pupils in a given group. If all the students from the upper and lower group answer an item correctly or if all fail to answer it correctly, the item has no validity, since in neither case does the item separate the good from the poor members of the class. An estimate of an item's discriminating power can be obtained by the formula :

$$\text{Discriminating Power} = \frac{U - L}{\frac{1}{2} N} \quad \ldots \quad \ldots \quad \ldots (2)$$

U = number of pupils from the upper group, who got the item right.
L = number from the lower group, answering correctly.
N = number of pupils in the two groups U and L.

Applying this formula to the item analysis data in Fig. 8.1, the index of discriminating power (D) would be :

$$D = \frac{15 - 7}{\frac{1}{2}(40)} = 0.40$$

If all the 20 pupils in the high group answer an item correctly and none from the low group gets it right, the discriminating power of the item is maximum and the index would be

$$D = \frac{20 - 0}{\frac{1}{2}(40)} = 1.00$$

The lowest validity index would be in an item which is answered correctly by an equal number of pupils say 16, in both the groups.

$$D = \frac{16 - 16}{\frac{1}{2}(40)} = 0.00$$

There may be a few items of negative validity, that is, an item wher more pupils in the lower group than in the upper group get the item right. Items having zero or negative validity must be revised or may be discarded.

(d) *Difficulty Value and Discriminating Power* : In a class of 20 pupils ten pupils get a score of zero and ten pupils get a score of 100 in a test in Mathematics, having 100 items. In this case the test has a poor discriminating power. The test cannot discriminate between the ten pupils who have scored zero. Similarly it fails to discriminate between the remaining ten pupils.

The basic function of all educational measurement is to place individuals along a defined scale in accordance with differences in their achievement. Such a function implies high discriminating power. An item must not be so easy as to be passed by every member of the group, nor so difficult as to be failed by every member. In neither case, does the item make any contribution to the discrimination which the test is to make between different individuals. As far as possible, no one should make a zero score in an achievement test; zero score pupils are unmeasured. A score of zero does not mean absence of any ability. Similarly no one should be able to make a perfect score. Perfect score pupils are unmeasured just as zero score pupils are. An item that does not in a very direct way serve to differentiate between pupils at various levels of achievement, has no place in a test since it adds only useless dead weight. A few very easy items may however be included at the beginning of a test in order to arouse confidence in the pupils.

Common practice in test construction is to attempt to prepare items covering a wide range of difficulty, though in general the test as a whole should have about 50 percent difficulty value for the average pupil. In the case of speed tests however, the difficulty level should be fairly low and approximately the same for all items. In diagnostic test, difficulty level is not an important aspect, since the purpose is not to rank the individuals according to achievement, but to locate the weak areas.

(e) *Effectiveness of Distractors* : How well each distractor is operating can be determined by inspection and there is no need to calculate an index of effectiveness. In general, a good distractor is one that attracts more pupils from the lower group than the upper group. Information concerning the efficacy of the distractors can be obtained by tallying the responses of the upper and the lower groups to each distractor. Let us consider three more items from a test in geography.

Geography

Example 1. Total no. of students = 80

Objective : Application

Specification : Cause and effect relationship

Item : People in Cochin need light clothing all the year round because

(a) they do not like woollen clothes
(b) the climate is sultry
(c) cotton is cheap in Kerala
(d) they are poor. [Answer : b]

Item Analysis Data

Alternatives	A	(B)	C	D	Omit	Total	Difficulty Value	Discriminating Power
Upper 25%	0	20	0	0	0	20	95%	0.10
Lower 25%	0	18	2	0	0	20		

Comments : This is a very easy item; its difficulty value is 95 percent. Its discriminating power is very low. Distractors A and D have not functioned at all, since none of the students chose them. When the distractors are illogical or obviously absurd, the examinee will have little difficulty in selecting the right option. When two distractors are ineffective, the item has only two options remaining and the chances of success by mere guessing are enhanced. This item should be revised by changing the distractors or it should be altogether rejected.

Example 2. *Objective* : Knowledge

Specification : Identifying relationship

Item : Ocean currents are the direct result of :

(a) the gravitational force of the earth
(b) changes in the atmospheric pressure
(c) difference in temperature of ocean water
(d) the attraction of the moon. [Answer : c]

Item Analysis Data

Alternatives	A	B	(C)	D	Omit	Total	Difficulty Value	Discriminating Power
Upper 25%	2	3	12	3	—	20	35%	·50
Lower 25%	5	7	2	4	2	20		

Comments : This item is fairly difficult, since only 14 out of the 40 pupils selected the right option (c). All the distractors A, B and D discriminate in the right direction; each is more attractive to the lower group. The validity of the item is fairly high. The item may be retained.

Example 3. *Objective* : Application

Specification : Establishing relationship

Item : Ocean currents will stop flowing if

(a) the earth stops rotating
(b) the rivers stop draining into the oceans
(c) the earth moves on its axis from north to south
(d) the oceans receive equal amounts of heat.

[Answer : d]

Item Analysis Data

Alternatives	A	B	C	(D)	Omit	Total	Difficulty Value	Discriminating Power
Upper 25%	8	3	4	4	1	20	30%	– 0.20
Lower 25%	4	5	3	8	—	20		

Comments : This is a poor item because more students from the lower group than from the upper group chose the correct answer. The item discriminates negatively. The question appears to be misinterpreted by the brighter pupils. Alternatives A and C are poor distractors since they attract more pupils from the upper group than from the lower groups. The item may be discarded, or it has to be radically revised.

Many teachers find it profitable to keep a file of items for future use. A good plan is to write the item on a card and to record the item analysis data as shown in Fig. 8.1, with necessary comments regarding the effectiveness of distractors. By indicating on the item card both the learning outcome and the subject matter, it is possible to file the cards under both headings. As the file grows, it becomes possible to select the majority of items from the file for any given test without repeating the items too frequently. Items which measure *understanding*, *applications* and *thinking skills* are difficult to construct and hence the test item file is especially valuable in these areas. A teacher can select the items from the file by taking into consideration the specifications of the items, their difficulty value and index of discrimination.

[4] Final Form of the Test

(a) *Selection of Items* : Items analysis of the preliminary form of the test provides the basis for revising the test and for preparing the final form to be administered. Items are selected by considering the difficulty and validity of each item. The poor ones are either selected or discarded. Selected items are then arranged in an ascending order of difficulty. For motivational purposes the first few items should be so easy that no body would miss them. The items should cover a wide range of difficulty.

(b) *Fixing the Time Limit* : Appropriate time limits are to be set for the final form of the test. This is done by considering the records of time taken by different individuals at the time of the preliminary try out. The time limit to be fixed will depend upon the purpose of the test, for example, if the test is primarily an achievement test, this time allotment is so set that most of the pupils tested have an opportunity to attempt all or nearly all the test items. Class-room achievement tests should generally have liberal time allowances. In tests for measuring proficiency in short-hand, typing and computational skills, speed is an important factor; hence the time limit should be so adjusted that even the best pupil does not have enough time to finish the test. On the other hand more time should be allowed for diagnostic tests than for achievement tests. Time allowances must always take into consideration the age and ability of the pupils, the type of items used and the complexity of the learning outcomes to be measured.

(c) *Directions for administering the Test* : Testing conditions should be as uniform as possible everytime the test is given to various groups. In planning for this, the test constructor should provide careful instructions for answering different types of questions and for recording answers. Directions to the pupils should be as clear, complete and concise as possible. The aim should be to make the instructions so clear that the weakest pupil in the group knows what he is expected to do, though he may not be able to do it. The pupil should be told how and where to mark the items, the time allowed and any reduction of errors to be made in scoring. These instructions should be printed on the test booklet. A few practice items should be given to make the pupils familiar with the testing procedure. In a standardized test, examiners have to follow the instructions rigidly in order to have uniformity of procedure. Instructions to examiners are to be given in the test manual.

(d) *Preparation of the Scoring Key* : Scoring should be done by a pre-determined method. This ensures objectivity in scoring. In case of objective type tests a scoring key is always prepared in advance and all scoring is done according to the key. The scoring procedure should be fairly simple. As a rule, the best procedure in scoring objective tests is to give one point of credit for each correct response. It is not necessary to give differential scoring according to estimated difficulty or importance. The test maker must decide how he is going to treat guessing in his scoring procedure. There has been an argument that guessing should be discouraged in order that the score can represent real attainment unaffected by lucky guesses. The usual correction formula based on the assumption that the person who does not know the answer, will make a random guess, is :

$$S = R - \frac{W}{n-1} \quad \ldots \quad \ldots \quad \ldots \quad (3)$$

where S = score corrected for guessing.
R = number of correct responses.
W = number of incorrect responses.
n = number of options presented for each item.

Note that items, omitted by a student, do not count in this formula for guessing. For a true-false item, the number of options i.e., n is 2; hence the formula would be :

$$S = R - \frac{W}{2-1} = R - W \quad \ldots \quad \ldots \quad (4)$$

For a multiple choice item with 4 options, the formula would be :

$$S = R - \frac{W}{4-1} = R - \frac{W}{3}$$

In a true-false test of 60 items, Jayant got 30 right, 10 wrong and omitted the rest, while Tanuja got 30 right and 30 wrong.

Their scores would be :

Jayant .... $S = 30 - 10 = 20$

Tanuja .... $S = 30 - 30 = 0$

This formula assumes that when a pupil does not know the answer to an item, he guesses blindly among all alternatives and he selects the correct answer on the basis of chance alone. If Tanuja is wholly ignorant, she would guess at each item and her score would be 30 right and 30 wrong. Hence the score representing her real achievement would be $30 - 30 = 0$. If on the other hand, she answers 30 items correctly and omits the other 30, her score would be $30 - 0 = 30$. In a multiple choice test with 60 items, in which each item has 5 options, the above scores would be

Jayant : $$S = 30 - \frac{10}{5-1} = 30 - 2{\cdot}5 = 27{\cdot}5$$

Tanuja : $$S = 30 - \frac{30}{5-1} = 30 - 7{\cdot}5 = 22{\cdot}5$$

If no items are omitted by the students, or if they all omit the same number of items, their relative scores will be the same, regardless of whether or not this formula is used. In power tests when students get sufficient time to attempt all items, the formula may not be used. It is only in speed tests where some students omit a fairly large number of items while others omit a few, that the correction formula is used. In speed tests, pupils could increase their scores appreciably, by blindly guessing at the remaining untried items, just before the testing period ended. Whenever the correction formula is to be used, pupils should be informed, so that they will not indulge in random guessing.

[5] Administering the Final Form of the Test

(a) *Establishing Norms* : We have considered so far two important factors essential to standardization namely, uniformity of administration and of scoring. Two more conditions essential to standardization are : establishing adequate norms and testing the reliability and validity of the test.

Norms provide the user of a standardized test with the basis for a practical interpretation and application of the results. Psychological tests have no predetermined standards of passing or failing. An individual's score can be interpreted only by comparing it with the scores obtained by others. Hence in the process of standardization, it must be administered to a large, representative sample of the type of subjects for whom it is designed. This group is known as the *standardization* sample. The problem of selecting such an appropriate group is a very complex one. Norms obtained from a sample of children from urban areas, will not be applicable to those from rural areas. The factors that affect the scores include age, sex, geographical area, cultural background, socio-economic level, rural or urban environment, education of parents, experience and a host of other circumstances. The test maker must therefore consider all these factors while choosing a representative sample for administering the final form of the test.

Norms, to be useful, must be based on a representative sample of the population. A test administered to 5000 pupils only from Bombay or from Madras would not give results representative of pupils from other areas. Merely increasing the number from 5000 to 10,000 pupils in Bombay or in Madras would not be helpful. Such a sample is a biased one. Statisticians recognize fully that the amassing of more and more cases can never compensate for inadequacies of the theory and quality of sampling. Utmost care has therefore to be taken while selecting the sample.

We shall discuss the interpretation of scores with the help of norms in the next chapter.

(b) *Establishing Validity and Reliability* : All the steps discussed above, are a prerequisite for the process of standardization. After the norms have been derived on the basis of the final results, it remains for the test maker to obtain full evidence concerning the validity and reliability of the test. No matter how carefully the test is prepared, its merits should be established and not merely assumed. The test maker thus evaluates the test through this procedure. It serves as a final check regarding the efficacy of the test. We shall discuss these characteristics of a test in the next chapter.

If all these stages of standardization are thoroughly and properly executed, the test will produce satisfactory results. The studies of validity should be continued by trying the test with different types of groups. Such later studies will enhance the utility of the test and it can then be used for various purposes.

(c) *Manual of the Test* : Every standardized test should be accompanied by the test manual. The purpose of the manual is to explain in detail what the test is supposed to measure, how it was constructed, how it should be administered and scored and how the results should be interpreted and used. It should explain the nature of the sample selected, the number of cases in the sample and the procedure of obtaining the norms. The manual should openly display the weaknesses as well as the strengths of the test and should provide examples of ways in which the test can be used to advantage and warnings concerning limitations and possible misuses of the results.

The test so prepared should display some qualities. What are these qualities ? How can they be measured ? We shall discuss this aspect in the next chapter.

FOR FURTHER STUDY

1. Explain the various steps involved in the process of standardization of a test.
2. What is meant by item-analysis ? What is its importance ? Illustrate the procedure with suitable examples.
3. How does the purpose of a test affect
    (i) the content of the test ?
    (ii) the time limit to be allowed ?
    (iii) the difficulty of items ?
4. Explain the terms : Discriminating power and difficulty value of a test item. How do you determine these two indices ?
5. *Answer in brief the following* :
    (i) In what respects does a standardized test differ from the routine class-room test ?
    (ii) Explain the statement : 'Zero score pupils are unmeasured.'

(iii) What precautions should be taken in choosing the standardization sample ?

(iv) Why should a standardized test be accompanied by a manual ?

(v) Why should a test be so difficult that no pupil can make a perfect score and yet so easy that no pupil will make a zero score ?

6. (a) Explain the correction formula used in scoring objective type of test items.

(b) Vinayak marked 41 out of 50 items correctly and failed in the remaining items. Each question has four choices. What would be his score after correction for guessing ?

(c) On an 80 item multiple choice test with five choices each, Nanda gets 56 items right, 20 items wrong and she has omitted the rest. What will be her score after correcting the score for guessing ?

7. *Explain the following statements giving reasons :*

(i) Difficulty value of items is not an important aspect in a diagnostic test.

(ii) The problem of selecting the standardization sample is a complex one.

(iii) Merits of a test should be established and not merely assumed.

(iv) Items should be arranged in an ascending order of difficulty.

(v) A mere collection of good items does not necessarily make a good test.

Answers :

6 : (b) 38 (c) 51

* * *

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

9

ESSENTIAL QUALITIES OF MEASURING INSTRUMENTS

Any measuring instrument whatsoever must fulfil some conditions if it is to be of any service. This is true of foot rules, balances, thermometers, speedometers and also of evaluation tools. A test or evaluative technique is judged for its adequacy, efficiency, and consistency as a measuring device on the basis of commonly accepted qualities. These qualities are *validity, reliability, objectivity, adequacy, usability,* and *discrimination.* The term validity refers to an instrument's truthfulness, reliability to its consistency and usability to its practicality.

(i) A good test must actually measure what it is supposed to measure (validity).
(ii) It must do this accurately and consistently (reliability).
(iii) It must be fair to the students (objectivity).
(iv) It must be long enough to do the job (adequacy).
(v) It must be easy to use (ease of administration and scoring).
(vi) It must pick out the good and the poor separately (discrimination).

All these factors are interdependent. They affect each other. Validity is affected by all the other five characteristics of the test.

1. Validity

It may be defined as the accuracy with which a test measures whatever it is supposed to measure. A test is valid, if it meets the purpose for which it was designed. Validity is a matter of degree. It does not exist on an *all-or-none* basis. We may consider validity in terms of categories such as high, moderate or low validity. The results of an arithmetic test may have a high degree of validity for indicating computational skill, a low degree of validity for indicating arithmetical reasoning and no validity for indicating success in art or music. Thus when appraising or describing validity, it is necessary to consider the use to be made of the results.

Furthermore validity is specific; a test may be valid for one purpose but not valid for another. An ordinary foot-rule is useful and therefore valid for measuring lengths of straight lines such as 4·6″ or 4·7″, but it is not a valid instrument for discrimination between lengths such as 4·66″ and 4·67″ or 4·86″ and 7·89″ simply because the foot-rule does not have these finer divisions on it. A scientist would not rely on a foot-rule for this purpose. He will have to use another instrument designed specially for this purpose. Similarly no body would use a foot-rule to measure the distance between Bombay and Mysore. A doctor's thermometer is valid for measuring body temperatures but not valid for measuring temperature of hot water ranging between 32°F and 212°F. Similarly nobody would use a laboratory thermometer for measuring atmospheric pressure. A speedometer valid for measuring the speed of a car is not valid for measuring the speed of a space ship or of a jet plane. An achievement test in geography for fifth grade students in Gujarat may not be necessarily valid for fifth grade students in other States, since the content of geography course may be different from one State to another.

Types of Validity

Four types of validity have been identified and are now commonly used in educational and psychological measurement. They are : *content validity*, *predictive validity*, *concurrent validity* and *construct validity*. These types are applicable to all types of evaluation tools.

What are the factors that influence the validity of a test ?

(1) *Unclear Directions* : When the directions do not clearly indicate to the pupil how to respond to the items, how to record the items etc., the answers will tend to reduce the validity.

(2) *Reading Vocabulary and Sentence Construction too Difficult* : Sometimes the vocabulary and the sentence structure used in the question is very complicated for the pupils taking the

test. The students might be knowing the answer but they fail to write it simply because they do not understand the question. Thus the geography or science test which uses such difficult language becomes a *reading comprehension* test and it does not measure what it intends to measure.

(3) *Medium of Expression* : When the medium of instruction was English at the high school and intermediate college stage, many students though they knew the subject matter, used to fail in such subjects as history or geography or science, simply because they failed to express the matter through English. Thus the examination instead of measuring the outcomes in history or geography measured primarily the knowledge of English. Such an examination does not measure what it intends to measure and the validity is lowered.

(4) *Difficulty Level of Items* : Items which are too easy or too difficult will not provide reliable discriminations among pupils and will therefore lower the validity of the test. If in an achievement test in mathematics 20 percent of the students get a zero score, it means the test is not able to discriminate among these 20 percent pupils and hence it has low validity.

(5) *Poorly Constructed Test Items* : Test items which unintentionally provide clues to the answer tend to measure primarily a pupil's alertness. Thus an alert pupil get the benefit more than the one who may not be so alert, though he may be knowing the subject matter equally well.

(6) *Use of Inappropriate Items* : Class-room tests are generally so constructed that they measure the *knowledge* objective. These questions give undue importance to the ability of pupils to memorize factual information. Other objectives such as applications, thinking, and skills are not measured through these routine type of questions, which ultimately invalidate the results. If you use a written test for judging the laboratory skills of students, it will have little validity.

(7) *Inappropriate Time Limit* : In a speed test, where the time factor is important for deciding the performance, if ample time is given, it would invalidate the results. On the other hand, if time limit given in an achievement test is too inadequate, the fast writer will get an advantage over the slow writer. The test instead of measuring achievement will measure the speed of writing.

(8) *Influence of Extraneous Factors* : Generally in essay type tests the examiner is greatly influenced by such factors as style of expression, mechanics of grammar, length of the answer, handwriting and the method of organization. A student having a good handwriting consistently gets a higher score in an essay examination. If the higher score is partly due to good handwriting, the score is not telling the whole truth about the achievement, hence the test is not valid. All these factors lower the validity of the test. The practice of deducting marks for lack of neatness or for poor handwriting is not a sound one. These traits may be assessed separately.

(9) *Inadequate Coverage* : When a test covers only 5 or 6 topics out of a total number of 20 topics, the validity of the test is called into question because of inadequate sampling. It fails to measure adequately what it is supposed to measure. This happens mostly in essay type questions.

(10) *Halo-Effect* : If a teacher has formed a poor impression about Leela and good impression about Meena, she is likely to carry over this impression while assessing the answer books of these girls. In rating scales, generally we find this effect. The halo-effect lowers the validity of the rating scale and of other tests.

(11) *Questions in the Form of Quiz* : In essay type examinations students sometimes do not understand the exact implication of the questions. They just guess and write the answers. The test thus measures *the ability of the students to guess* or to solve *a quiz*. Consider the following questions that have appeared in some of the University examinations :

(i) Discuss : "The personality of the teacher means everything in education."

(ii) Explain : "We learn to skate in summer and swim in winter."

(iii) Explain : "Vocational guidance means helping Johnny to see through himself and to see himself through."

(iv) Comment : Heurism must guide our teaching but it must not be allowed to control it."

If such statements are taken from the next-book we can expect the students to know their meaning. In that case validity is not affected. But if the statement is picked out from some other book which the students might not have read, the validity of the test is lowered. It will be seen from the above discussion that *the validity is lowered whenever the test measures something else than what is intended*. Hence while preparing a test, we have to be always on guard regarding the purpose of the test. When we lose sight of the purpose, other extraneous factors may creep in. and will lower the validity of the examination.

2. Reliability

The second important quality of a measuring instrument is its *reliability* or *consistency*. A reliable test is one which measures accurately and consistently. If a reliable test is given two or three times to the same group, each person in the group should get approximately the same score on all occasions. If the relative standing of pupils differs when the test is administered twice to the same group, we may infer that the test is not reliable, since it is not giving consistent scores. No one would consider relying on a thermometer which gives readings varying between 96° and 104° F, for persons known to have normal temperature. Nor would any of us place confidence in measurements of length based on an elastic ruler. Reliability is a statistical concept. Little can be said about the reliability of a test by examining the test items themselves. (This is not so in estimating validity; content validity can be estimated by analysing the content of the test as discussed on page 124.) To determine the reliability of a test, it must be administered to a group of individuals under appropriate circumstances. Reliability can be expressed by a correlation coefficient which is called *reliability coefficient*. There are various methods for estimating the reliability.

(i) *Test-Retest Method* : The same test is given to the same group on two occasions and the correlation coefficient between the two sets of scores is determined.

(ii) *Parallel Forms Method* : Two parallel forms of the test are given to the same group on two occasions and correlation between the scores on the two forms is calculated.

(iii) *The Split Half Method* : The most widely used procedure for estimating reliability from single testing is the split-half method. The test is divided into two halves as equivalent in difficulty and in content. For each individual two scores of these two halves are obtained. Correlation between the halves is determined and is taken as a measure of the accuracy of the test.

What are the factors that influence the reliability of a test ?

A number of factors have been shown to affect reliability. If sound conclusions are to be drawn, these factors must be considered while interpreting reliability coefficients. We should not compare two tests for their accuracy by considering the face value of the reliability coefficients. Conditions under which these coefficients have been obtained, must also be considered.

(i) *Length of the Test* : Lengthening a test affects directly its reliability. A test of 50 items is more reliable than one of 20 items, which in turn is more reliable than a test of 2 items only. This is due to the fact that a longer test will provide a more adequate sample of the behaviour being measured and the scores are apt to be less influenced by chance factors such as guessing. In lengthening the test, it is assumed that items that are added are of the same quality as those already in the test. Lengthening a test by adding poorly constructed items will not increase its reliability.

(ii) *Objectivity of Scoring* : Objective type of tests in general produce more reliable results than the subjective or essay type tests. But objectivity should not be confused with reliability. It is possible for a test to be quite *unreliable* as a measuring instrument yet have high scoring *objectivity*. The true-false test can be objectively scored but it may be unreliable because of the element of chance in selecting the answer.

(iii) *Range of Talent* : A yardstick may be found reliable to differentiate ten lines varying in length from 30″ to 40″ but the same yardstick may be unreliable if the ten lines vary minutely from 30″ to 30·4″ This is true of an evaluation technique also. A test may be found reliable when applied to a group of 60 students with a wide range of ability. Suppose the scores range from 20 to 90. On a second trial, the relative standing of the students will most probably remain the same; if however, the range of ability of the group is very narrow as for example, the scores on the test range from 50 to 62, it is most likely that the relative ranking of these 60 students will change on a second administration of the test. That is, the test which is reliable for the first group may not be necessarily reliable for the second group.

(iv) *Difficulty of the Test* : Tests which are too easy or too difficult for the group, will tend to provide scores of low reliability. Because in both cases the variability of the scores is very small. Differences among individuals in such tests are small and tend to be unreliable.

(v) *Ambiguous Wording of Questions* : When questions are interpreted in different ways at different times by the same students, the test would not give consistent results. The interpretation differs because of ambiguity in the wording of the questions. The student is not certain about the meaning of words. This brings in guessing into the situation.

(vi) *Testing Conditions* : Sometimes conditions may change while administering and scoring the test. These changes make the scores unstable and thus produce unreliability in a test, for example, slight changes in the time limit or changes in instructions, fluctuations in interest and attention of the pupils, shifts in emotional attitude all these factors may adversely affect the reliability of a test.

(vii) *Use of Optional Questions* : Giving a choice of questions reduces the common base on which different individuals may be compared. The choice adds to the unreliability of test scores. Students may not attempt the same items on a second administration of the test. An individual's score will not be consistent from one administration to another.

What is the relation between reliability and validity ?

Test makers who have no real evidence about the validity of their tests emphasize the high reliability of the test. It is therefore important to understand the relationship which exists between the validity and reliability of a measuring device. The relationship is expressed as follows :

A reliable test may not be necessarily valid; but a test cannot be valid unless it is reliable. Reliability is therefore a necessary prerequisite to validity but does not in itself guarantee validity. In fact it is possible for a measure to be extremely reliable and yet possess no validity whatsoever for a particular purpose. For example, the circumference of the head can be measured with great precision; however, this measure has been shown to have no significant correlation with intelligence and thus is not a valid measure for the purpose of estimating intelligence. This does not however mean that it is not a valid measure for any other purpose. After all it is a perfectly valid measure for determining the size of a hat, a person should buy. A well calibrated yardstick gives accurate results. It is therefore reliable. It may be highly valid for measuring the height of school children but of little validity for measuring the distance between Bombay and Baroda or for determining the weight of apples in a basket. Similarly a good test of arithmetic skill (which is reliable), would be a valid measure of pupil proficiency in skills such as addition and subtraction, but of low validity for problem solving ability in arithmetic and of even lower validity for the solution of geometrical problems.

Another aspect of this relationship is : An inaccurate measuring instrument, which gives inconsistent results, cannot possibly provide truthful information. A badly worn out yardstick could not be expected to measure anything accurately; it could not therefore be considered valid for any measurement purpose. Similarly a poorly constructed and hence an inaccurate test of arithmetic skills could not be used in obtaining a valid measure of any type of skill in arithmetic.

A measuring device with a reliability of 0.00 is reflecting nothing but chance factors. It does not correlate with itself and cannot correlate with anything else. Validity is always the first quality to be sought in a test and reliability is then a valuable auxiliary. An ideal test tells the truth consistently. The two concepts *reliability* and *validity* refer to different aspects of what is essentially the same thing, namely test *efficiency*. Reliability refers to *trustworthiness* while validity refers to *truthfulness.*

3. Objectivity

A test is objective when the examiner's personal judgment does not affect the scoring. An objective test is one for which the items can readily be scored as right or wrong. Such item types as multiple-choice, matching, true-false are objective, but the essay examination is subjective rather than objective.

Objectivity is a prerequisite to reliability in a test; subjective judgments are inaccurate and hence unreliable. Objective judgments are accurate and hence tend to be reliable. In essay type examination as well as in rating scales, the results depend to a large extent upon the person doing the scoring. Different persons get different results; such inconsistency in scoring has an adverse effect on the reliability of the measures. The solution is not to use only objective tests and abandon all subjective methods of evaluation. A more desirable solution is to make the evaluation tool as objective as possible. In the case of easy tests, for example, objectivity can be increased by careful phrasing of the questions and by using a standard set of rules for scoring. Another aspect of objectivity pertains to the interpretation of individual test items by the students. Well consructed test items should lend themselves to one and only one interpretation by students. Objectivity can be determined by finding the coefficient of correlation between scores, assigned to a group of papers by the same examiner on two occasions. It is called the *coefficient* of *objectivity.*

4. Adequacy

The careful test maker never assumes that the instrument he has constructed is capable of measuring all of the factual knowledges or skills that a pupil has acquired in a school course. Just as a grain-dealer samples a cart-load of wheat by taking samples from different bushels in the cart and grading the samples in order to obtain a measure of quality for the whole cart-load, a test constructor measures the educational attainments of pupils by constructing test items that represent widely the types of pupil outcomes expected and accepts the scores as representative of the pupils' achievement for the entire area sampled by the test items.

Adequacy is a prerequisite to the reliability as well as to validity of a test. Two apples from a basket full of apples might well constitute a very poor sample of the contents of the basket. Similarly a two item test in arithmetic or history will not be representative of the course content, hence it would lack in validity. Element of chance increases because the test is too small, thus it lacks in reliability. Hence a test must fulfil this criterion of adequacy as a prerequisite for reliability as well as for validity.

5. Usability

There is no doubt that the ultimate criteria of the value of a test are its validity and its reliability. But the most valid test in the world will do nothing to improve the school programme, if it is not used. *Usability* or *practicality* is an important criterion for assessing the value of a test. Usability of a test depends upon a number of factors. Among these are : ease of administration, ease of scoring, ease of interpretation and use of scores, low cost, satisfactory format and so on.

(i) *Ease of Administration* : Directions for administering the test should be simple and clear, the subtests should be relatively few and the timing of the test should not be too difficult for adjustment.

(ii) *Ease of Scoring* : One of the most troublesome aspects of a school testing programme has been the scoring of tests. If algebraic manipulations are required to get the scores or if the original raw scores require to be converted into complicated derived scores, teachers will avoid to use such a test.

(iii) *Ease of Interpretation* : In the final analysis, the success or failure of a testing programme is determined by the use made of test results. While selecting a test, attention should be directed toward the ease with which raw scores can be converted into meaningful derived scores and toward the clarity of suggestions for applying the test results.

(iv) *Economy* : Economy is certainly not one of the major criteria of a good test but it is a factor that must be considered. Real economy in testing will not be achieved by the indiscriminate use of cheap tests, but it is equally true that users will avoid buying costly instruments and tests. Economy also refers to the time required for administering and scoring the test. Use of separate answer sheets and machine scoring will reduce the cost appreciably.

(v) *Availability of Equivalent Forms* : Equivalent forms of a test measure the same aspects of behaviour by using test items which are alike in content, level of difficulty, and other characteristics. Thus one form of the test can substitute the other. Equivalent or parallel forms are especially valuable when these are used for research purposes or for measuring the effects of teaching a particular course.

6. Discriminating Power

A test discriminates well when it is constructed in such a manner that it will detect or measure small differences in achievement and when it picks out the good from the poor. This is an absolute essential if the test is to be used reliably for ranking students on the basis of achievement. This quality of a test directly affects its validity. This point has been already discussed under *Item Analysis* in the previous chapter.

Interpretation of Test Scores

We have discussed above the essential qualities of measuring instruments. However good an instrument may be, its use cannot be justified, if its results remain unintelligible. If the result is misinterpreted or misapplied or not applied at all, the testing programme will be of little value. A test is a delicate instrument which must be cautiously used. A test will yield only a score and not the conclusion to be drawn from that score. Before considering the interpretation of test scores, we must know the limitations of educational measures.

Comparison of Physical and Educational Measurement : Test interpretation would be greatly simplified if we could express test scores on scales like those used in physical measurement. We know for example, that 5 ft. means the same height whether we talk about the height of a boy or a door or a wall. Similarly we can say that Suresh has a weight of 160 lbs. and weighs exactly *twice* as much as Paul who weights 80 lbs. or that a period of 10 minutes is exactly *one fourth* of a period of 40 minutes or a weight of 10 gms. is *equal to* two weights of 3 gms. and 7 gms. This ability to compare measurements from one situation to another and to speak in terms of 'twice as much as' or 'one fourth of' is made possible by the fact that these physical measures are based on scales which have a *true zero point* and which have *equal units.* The true zero point (e.g. the point at which there is no height at all or no weight at all) indicates precisely where measurement begins and the equal units (e.g. feet, pounds, minutes etc.) provide uniform meaning from one situation to another and from one part of the scale to another. Ten pounds indicates the same weight to the doctor, the grocer, the farmer and the housewife. Also, the difference between 20 and 28 lbs. represents exactly the same difference as that between 60 and 68 lbs.

Unfortunately, the properties of physical measuring scales with which we are so familiar are generally lacking in educational measurement. A pupil who receives a score of *zero* in an arithmetic test does not have *zero knowledge* of arithmetic. There are probably a large number of simple questions he could answer which were not included in the test. *A true zero point* in achievement where there is no achievement at all, cannot be clearly established. Even if it could, it would be impractical to start from that point each time we tested. What we do in actual practice is to assume a certain amount of basic knowledge and measure from there. This arbitrary starting point, however, prevents us from saying that a score of zero indicates 'no achievement at all' or that 'a score of 80 represents twice the achievement of a score of 40' or that 'a score of 60 represents an achievement equal to the achievement of two students with scores 12 and 48. Since we are not certain how far the zero score on our test is away from the true zero point (i.e., the point of no-achievement at all), test scores must be interpreted in relative rather than absolute terms. Test scores are not *absolute* but *relative* ones. We can speak of 'more' or 'less' of a given achievement but not 'twice as much' or 'half as much as' of something.

The interpretation of test results is additionally handicapped by the *inequality* of our units of measurement. Forty items correct on a simple arithmetic test does not have the same meaning as 40 items correct on a more difficult test in arithmetic; nor do these scores represent the same level of achievement as 40 items correct on a test in science or history. Similarly a change of score from 60 to 70 does not mean the same improvement as a change of score from 10 to 20.

Thus the yardstick used in educational measurement is a peculiar one. (1) It does not give absolute measures but only relative ones; it does not show the true zero point, meaning no achievement at all. (2) The yardstick is also not uniform throughout; it is like a rubber yardstick that has been stretched out at some points and squeezed in at others, i.e. it has no equal units. We can at the most say that Sheila is more intelligent or more confident or more obstinate than Lynda, but cannot use any ratio to show that Sheila is twice or thrice as intelligent or confident or obstinate as Lynda is. (3) The yardstick does not give scores which will have uniform meaning from one test to another. To overcome the lack of a definite frame of reference in educational measurement, various methods of expressing test scores have been devised.

Types of Test Scores and Norms

Raw Scores : If a pupil responds correctly to 40 items on an objective test, in which each correct item counts one point, his raw score will be 40. *Thus a raw score is simply the number of points received on a test when the test has been scored according to instructions.* A raw score provides a numerical summary of a pupil's test performance, but it is uninterpretable without further information. Getting 40 items correct in a language test for example, has no particular meaning in itself. We cannot determine whether it is high or low unless we have some basis for comparison. Our score of 40 would be high if there were 50 items in the test (80 percent correct) but low if there were 100 items i. e. 40 percent correct. Moreover since we have no absolute standard for judging the performance, the term 40 percent correct or 80 percent correct is meaningless. If the test is easy, getting 80 percent correct may not be even considered as a good performance.

Derived Scores : The use of raw scores is restricted in two important ways. First since a raw score is meaningless by itself it is difficult to interpret the score. To make a raw score of 40 interpretable, for example, it must be accompanied by information concerning the nature of the group tested and the distribution of scores obtained. Second, raw scores on different tests cannot be compared directly. If a pupil obtains a raw score of 50 on a language test and 56 on a mathematics test, we cannot definitely say that his performance in mathematics is better than in language. Getting a score of 50 in a language test may indicate superior achievement while a score of 56 in mathematics may be considered as average achievement. To compare achievement on different tests, we need a unit of measurement which has fairly uniform meaning from one test to another and from one part of the scale to another (like those in physical measurements). The problem is similar to that of the comparison of vulgar fractions: we cannot directly add or compare the factions such as :

$$\frac{2}{3}, \frac{3}{4}, \frac{5}{6}, \frac{6}{8}, \frac{7}{12}$$

First we have to find a common unit for all these fractions, that is a common denominator which is 24 in this case. Then we can express all the fractions in terms of that denominator. The fractions would then be :

$$\frac{16}{24}, \frac{18}{24}, \frac{20}{24}, \frac{18}{24}, \frac{14}{24}$$

These are now comparable because of a common unit. Similarly test scores must be changed so as to have a common unit or a common denominator. These scores are then called *derived scores.* These scores provide units which are approximately uniform. *A derived score is a numerical report of test performance in terms of the pupil's relative position in a clearly defined reference group.* Such a score merely describes an individual's position in some group. Converting raw scores into derived scores is simple with standardized tests. All we need to do is to consult the table of norms and select the derived score which corresponds to the pupil's raw score. In case of standardized tests, a derived score is a numerical description of a pupil's performance in terms of norms such as mental age, IQ, reading age etc.

Norms

Tables of norms included in test manuals merely present scores earned by pupils in clearly defined reference groups. Norms are levels of performance on a test attained by defined groups of pupils. Norms differ from standards in that they describe *what is* rather than *what should be*. They do not represent especially good or desirable performance but rather normal or typical performance. They should not be viewed as standards or goals to be achieved by other pupils. Norms are worked out for groups of individuals but a goal or a standard may be defined for each individual. There are different types of norms such as *age norms, grade-norms, percentiles, standard scores* and *quotients*.

1. Age Norms

For any trait that shows a progressive change with age such as height, weight, mental ability, reading ability, achievement in school etc., one can prepare a set of age norms. To obtain age norms we give a test to representative groups of pupils at various age levels, and determine the average of the scores obtained in each age group. Many achievement tests and most intelligence tests are provided with age norms. Thus we speak of a pupil's reading age, arithmetic age, spelling age and the like. Sheila who is ten years old, has a score of 65 on a reading test. This score is the average score of 12 year old pupils on this test, hence Sheila received a derived score of twelve or we may say that her reading age is twelve (Table 9.1). If Sheila's score on an arithmetic test corresponds to the average score of 8 year olds, her arithmetic age would be eight. In intelligence testing, age equivalents are called *mental ages* (MA). Jayant who is ten years old, has a mental age of eleven. It means that his performance on the intelligence test equals that of the 11 year old children. In achievement tests the term *educational age* (EA) is used to show the achievement over broad areas of subject matter. It shows the level of educational growth.

Merits: (i) Age norms are easy to understand. This is especially true for measures highly correlated with age such as mental ability, reading ability and various skills taught in the elementary school, which have fairly consistent growth patterns. (ii) This frame of reference is a stable one and is convenient for use.

Limitations of Age Norms : (i) Age units are characteristically unequal. Mental ability develops rapidly up to the age of 12, then development slows down and after 15, the yearly increase is not large enough to form the basis of a scale. The mental age unit does not remain constant from age to age but tends to shrink with advancing years. A year's growth from age 3 to age 4 is not the same as the growth from age 13 to age 14. (ii) Age norms are not useful to express the development of personality traits such as sociability or extroversion-introversion, which do not have fairly consistent growth pattern or which are not highly correlated with age. (iii) Mental age remains a useful concept at the high school level but its meaning is not so clear for *above average* pupils. Since mental growth slows down at approximately age sixteen, mental ages beyond this must be estimated by extra-polation. (iv) Two years' acceleration or retardation may have different meanings depending on the age-level at which it occurs. Two years' retardation at the age of 4 is more serious than one at the age of 8 or 10. At the earlier age it represents a larger deviation from the norm than at the later age. (v) The mental age concept is actually inapplicable for chronological age levels above fifteen years. Mental age does not increase with chronological age throughout an individual's life time.

2. Grade Norms

Grade norms are widely used with standardized achievement tests especially at the elementary school level. They are based on the average scores earned by pupils in each of a series of grades and are interpreted in terms of grade equivalents. For example, if pupils in the standardization group who are beginning the 6th grade, earn an average score of 57 on a reading test, the score is assigned a grade equivalent of 6·0. Intermediate values such as 6·4, 6·5, or 6·6 may be found by interpolation or by testing the students at different times during the school year. Thus 6·0 in the second column of Table 9.1 means the norm for the beginning of the sixth grade, 6·5 is the norm for the middle of the sixth grade while 6·10 means the norm for the end of the sixth grade which is the same as the norm for the beginning of the seventh grade. As the school year is of ten months, decimal system is used to show the grade and months.

Table 9.1 : Age and grade norms for a reading test.

Raw Score	Grade Placement	Reading Age
57	6.0	11—0
58	6.1	11—3
59	6.2	11—5
60	6.3	11—6
62	6.5	11—8
64	6.6	11—10
65	6.8	12—0
66	6.9	12—2
67	7.0	12—3
68	7.2	12—6
70	7.3	12—7
71	7.5	12—10
72	7.6	12—11
74	7.7	13—0
75	7.9	13—3

3. Quotients

(*a*) *Intelligence Quotient* (*IQ*) : In the early days of mental testing, after age norms had been used for a few years, a need was felt to convert the age score into an index that could express rate of progress. *Mental age, as we have already seen shows the level of mental development but does not show the rate of development.* A 4-year old child has a mental age of 6-years; he is better than average, but how much better? An 8-year old has a mental age of 10 years. We see that both have an acceleration of 2 years. Are they maturing at the same rate? We cannot answer these questions regardless of the chronological age of the student. Therefore an index was needed that would take account of both the chronological and mental ages. The first man to suggest a measure which unlike mental age would permit a uniform interpretation regardless of the age of the subject was William Stern, a German psychologist. He called his measure *Mental Ratio*, (MA/CA). Four years later Stanford University psychologist Lewis Terman multiplied Stern's mental ratio by 100 to remove the decimal and gave the world a new term 'Intelligence Quotient' (IQ). This was first employed in the 1916 Stanford Revision of Binet Scale.

$$IQ = \frac{MA}{CA} \times 100$$

The IQ in the above two cases would be :

$$(1)\ IQ = \frac{6}{4} \times 100 = 150 \qquad (2)\ IQ = \frac{10}{8} \times 100 = 125$$

The younger of the two is more intelligent, his rate of development is faster than that of the other.

Limitations of IQ : (i) IQ's on one test are not comparable to those obtained on another test. (ii) We have already seen that the mental age concept is essentially inapplicable to normal and superior adults. Hence IQ's in case of adults are only hypothetical and are open to doubt. (iii) IQ alone does not indicate the level of maturity of a child; it only tells the rate of maturity. Two children having an IQ of 120 may not be equally mature. We can say they are above average when compared with their own age group. A child of 8 and a child of 10, may have the same IQ but their level of maturity is different, which is indicated by their mental ages. (iv) In common with all units in which test scores are expressed, the IQ suffers from two limitations. The zero point is arbitrary rather than real and the various units are of unequal length. The difference between an IQ of 60 and 70 is not the same as the difference between an IQ of 110 and 120.

(*b*) *Educational Quotient* (*EQ*) : Similar quotients can be computed for achievement tests, reading tests or any other testing instrument that yields age norms. Thus the Educational Quotient (EQ), or the Reading Quotient (RQ) is the ratio between a child's educational age or reading age and his chronological age. Such a quotient is a measure of the rate of educational growth. As measures of the rate of progress, they may be useful in the interpretation of scores on achievement tests. However no elaborate scheme for the interpretation of these quotients has been worked out as is seen in the case of IQ's of 80, or 100, or 120 or 150.

$$\boxed{\begin{aligned} EQ &= \frac{EA}{CA} \times 100 \\ RQ &= \frac{RA}{CA} \times 100 \end{aligned}}$$

(*c*) *Accomplishment Quotient* (*AQ*) : An especially fallible and difficult to interpret ratio that has all but vanished under the onslaught of criticism is the *accomplishment quotient*. It represents an attempt to obtain a measure of whether a pupil is achieving the objectives of class instruction to the limits of his mental ability. Formula for this quotent :

$$AQ = 100 \times \frac{EA}{MA} \quad \text{or} \quad 100 \times \frac{EQ}{IQ}$$

If a pupil's achievement (EA or EQ) is in keeping with his ability to learn (MA or IQ), his AQ will be 100. Obviously the indication of an AQ below 100, should be that the child is not working to capacity and an AQ of more than 100 would be impossible, as one cannot normally achieve beyond one's capacity. A probable explanation is that the higher value is due to the inaccuracies in the tests used to determine EA and MA.

Limitations : This quotient, though popular at one time is now practically extinct. First of all, it tends to penalize the bright pupil who cannot work up to capacity as readily as the normal or the dull pupil because the instructional levels of most schools are geared to the average and inferior pupils. The curriculum does not have enough challenging material to interest and to motivate superior pupils. Secondly this quotient has low reliability because of compound errors; both the numerator and the denominator (EA and MA) do not give accurate results. Hence a ratio between two measures that are themselves not highly reliable (in this case the EA and MA) cannot be highly reliable because the quotient of two unreliable measures is less reliable than either of the measures. The use of IQ or EQ involves less risk than the use of AQ because the divisor CA used in computing IQ or EQ is almost wholly free from errors of measurement if obtained from a birth certificate. Thirdly the AQ will not be a valid measure unless the distributions of the two tests giving EA and MA, are approximately equal; for only then will a given EQ such as 120, be comparable with the corresponding IQ of 120. Unless the two distributions are the same, the two seemingly identical figures will have different meanings.

4. Percentile Norms

One of the most widely used and easily comprehended methods of describing test performance is that of percentile rank. A percentile rank (or percentile score) indicates a pupil's relative position in a group in terms of the percentage of pupils scoring below him. Thus if we consult a table of norms and find that Sujata's raw score of 64 in geography equals a percentile rank of 75, we know that Sujata has surpassed 75 percent of students from her group. 75 percent of the pupils from her class obtained a raw score lower than that of Sujata. The method of computing percentile ranks will be discussed in chapter 18. Percentiles higher than 50. indicate above average performance and those below 50 signify inferior performance. For uses and limitations of percentiles refer to chapter 18.

5. Standard Score Norms

The units based on percentiles are so unequal that we are led to look for some other unit that has the same meaning throughout its whole range of values. Standard score scales have been developed to serve this purpose. They have the advantage of providing equal units. The standard scores, also called *sigma* scores or *z* scores are expressed in terms of the mean and the standard deviation of typical groups. We shall discuss the method of converting raw scores into standard scores in chapter 18. There are numerous types of standard scores used in testing. They enable us to compare the standing of an individual on two or more tests (Fig. 9.1).

Deviation IQ : One of the widely used standard scores is the deviation IQ. Here the mean is set at 100 and the standard deviation at 16. Thus a pupil with an IQ of 84 has scored one standard deviation below the mean and a pupil with an IQ of 116 has scored one standard deviation above the mean. A given deviation IQ indicates the same relative ability at different ages. A given IQ of say 84, has the same meaning at all age levels. They can also be readily converted to percentile ranks and to other types of standard scores as is shown in Fig. 9.1.

Many test authors are using the deviation IQ's instead of the usual ratio IQ's. Standard scores in different forms (including the deviation IQ) appear to be gradually replacing other types of scores because of certain advantages they offer with regard to test construction and statistical treatment of data.

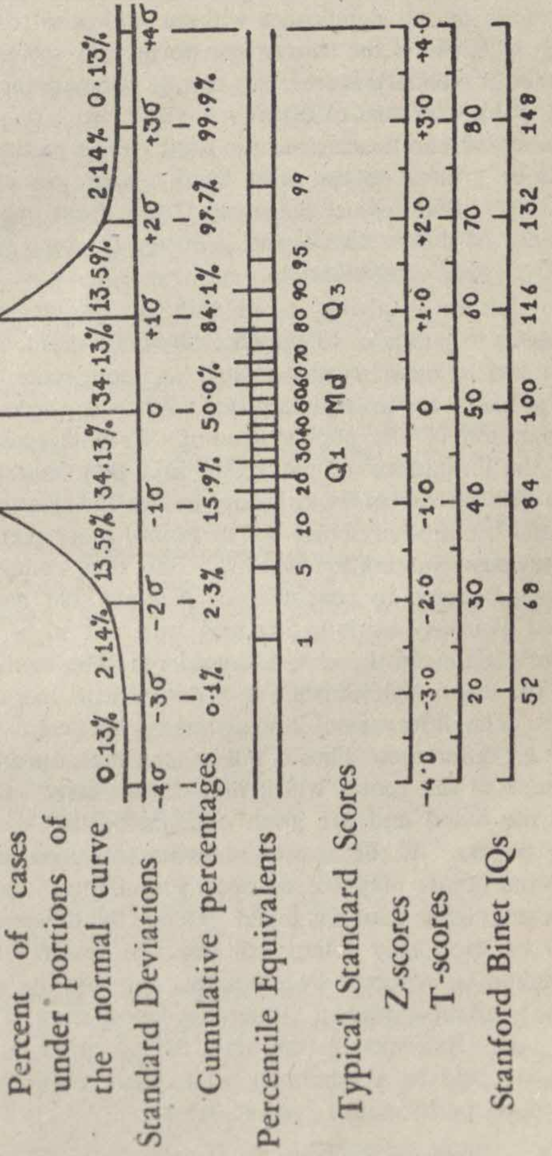


Fig. 9.1 : Methods of Expressing Test Scores.

Assigning Marks and Reporting Pupil Progress

As we noted earlier in this chapter, a raw score on a test has practically no significance without additional data for interpreting it. Even if the teacher transforms raw scores into percentile ranks or standard scores, they merely compare the performance of one pupil with that of others. A test score is only *an index of measurement* and measurement in itself has no particular meaning. *Marks or grades* on the other hand are *indices of evaluation;* they imply that value judgments have been made with the assistance of other criteria and in terms of some type of value, objective, goal or standard.

A teacher is always faced with the problem of reporting test results to pupils or to parents. She must decide *how much to report* and *in what way to report.* A mere score is not needed as it is likely to be misinterpreted. What is needed is a sound interpretation of the pupil's standing. Parents generally do not know the limitations of test scores and they misinterpret them. Sujata gets a score of 64 in history in the 2nd quarterly examination, and her mother scolds her by pointing out that her score in the previous examination was 68. She thus compares the raw scores and jumps to conclusions. Actually her performance in the 2nd quarterly examination may turn out to be better if her relative standing in the class is considered. The mother gets upset when she learns that Sujata's next door friend has secured 66 in history. The difference of 2 points makes her feel that it is a very significant difference. Thus it will be seen that a precision is being attributed to test scores which they do not have. Differences in scores are noted and are given significance that they do not in reality possess. If the answer-books are revalued, the difference of 2 score points may not necessarily remain. Sujata may even get a score higher than her friend. Hence the differences in scores should be cautiously interpreted and the results of class-room tests should be reported by statements showing the general level, such as : 'about average', 'somewhat below most of those in his group' etc. Presented in this way, the potential misuses of test results are held to a minimum, while their values in communicating pupil performance are enhanced.

What is a Mark or a Grade ?

A mark or grade in a school or college is in the last analysis a judgment of the student by the teacher. The marking system most frequently used is based on five points; these points are usually given in letter grades A, B. C, D and F. The major shortcoming of grades or marks is that few of them are defined meaningfully. The same marks often mean different things to different teachers even in the same school. It is therefore important to define each mark or grade to facilitate uniformity in understanding. The following definitions of grades may prove useful.

Significance of Various Grades

A. Excellent : this student is outstanding: he has achieved all the major goals of instruction; usually 5 to 10 percent of the pupils in a typical class receive this grade. His level of achievement is considerably above the minimum required for doing advanced work in the same field.

B. Highly satisfactory : the work of the student is generally better than average; his level of achievement is somewhat above the minimum, required for doing advanced work in the same field. Usually about 20 to 30 percent of the pupils in a typical class receive this grade.

C. Satisfactory : the student's work is quite acceptable but not outstanding. His level of achievement barely meets the minimum required for more advanced work in the same field. Usually about 35 to 40 percent of the students in a typical class receive this grade.

D. Poor : the work of the student is noticeably weak. He has achieved only a few of the major goals of instruction. He would require considerable remedial work and individual assistance. Usually about 20 to 30 percent of the pupils in a typical class receive this grade.

F. Unsatisfactory : this student does not meet the minimum requirements of the class. He has achieved none of the major goals. He has failed to accomplish the minimum essential for continued progress in the same field. Usually about 5 percent of pupils in a typical class receive this grade.

The percentages given above are likely to vary from one school to another. The marks will also depend upon the standards expected by the teachers. The standard is a shadowy inner picture which the teacher carries with her of the performance of her students. The grades defined above may be printed on the report card so that the parents try to interpret the gradation of their ward as per the definition.

The S. S. C. Examination Board (M.S.) Poona, uses two grading systems for two groups of subjects as mentioned on next page :

	Poor	Fair	Good	Very Good	Excellent
(i) Science, mathematics, and classical languages.	0–34	35–49	50–74	75–89	90 & above
(ii) Languages; history, geography, commerce, economics	0–34	35–49	50–64	65–74	75 & above

Many of the schools simply write the scores in the report cards which do not convey the right meaning to the parents. Hence along with the scores, following items may be included in the report card :

(i) Average score of the class in each subject
(ii) Rank of the student in percentile form
(iii) Gradation such as A, B, C, D or F with its definition
(iv) Attitude of the student and efforts made by him

Modified Report Cards

If the purpose of the report card is to transmit information about the child and if this information is to be interpreted correctly and used intelligently by parents, the report card must give adequate information in a meaningful manner. A sample report card is illustrated in Fig. 9.2. Scores on various tests should be shown separately as usual. These should be supplemented by information as given in Fig. 9.2.

1. Comments regarding the pupil's performance in various school subjects should be given by the teacher every quarter. A checklist is provided against each subject. (Fig. 9.2)

2. Learning outcomes in terms of interests, attitudes and habits must be also taken cognisance of in the report card. A checklist showing the various social habits, study habits and health habits is also given in Fig. 9.2.

3. Teacher should also give some comments regarding the achievement of the pupil in general and also the efforts put in by him.

It will thus be clear that mere scores on various examinations or mere letter grades such as A, B, C, D, F etc., should not be

PROGRESS REPORT CARD

Name :

Standard :

Progress in School Subjects

	Quarter I	II	III	IV
Reading				
1 Understands what he reads.	✓			
2 Works out the meaning and use of new words.	✓			
3 Reads well to others.	✓			
4 Reads independently for pleasure.	✗			
Mother Tongue				
1 Speaks clearly and distinctly.	✓			
2 Expresses effectively his ideas orally.	✓			
3 Expresses effectively in writing.	✗			
4 Writes legibly.	✗			
5 Listens carefully.	✓			
Arithmetic				
1 Is accurate in work.	✗			
2 Works at a satisfactory rate.	✗			
3 Solves problems involving reasoning.	✗			
Physical Education				
1 Participates actively in games.	✓			
2 Is developing muscular co-ordination.	✓			
Music				
1 Participates in singing.	✓			
2 Shows appreciation of music.	✓			
Art				
1 Is developing skills in expression.	✗			
Science				
1 Observes carefully.	✗			
2 Is precise in his statements.	✗			

Letter Grades

A—Excellent

B—Highly satisfactory or very good.

C—Average or Creditable.

D—Below Average or Poor

F—Unsatisfactory or failing

The teacher should use the above letter grades for denoting the total progress in school subjects as well as in various habits. She should offer her comments regarding the total achievement of the pupil and the efforts put in by him.

Tick marks may be made while checking periodically the various items given in this card.

Progress in Habits

	Quarter I	II	III	IV
Social Habits				
1. Uses self-control	✓			
2. Is observant of the rights and feelings of others	✓			
3. Cooperates with others in work and play.	✓			
4. Shows willingness to help others.	✓			
5. Accepts responsibilities.	✗			
6. Makes a neat personal appearance.	✗			
Study Habits				
1. Completes assignments in a reasonable time.	✗			
2. Participates in group discussion.	✓			
3. Follows directions.	✓			
4. Is neat and orderly in work.	✗			
5. Uses time wisely.	✗			
6. Shows independent judgment.	✗			
Health Habits				
1. Practises good health habits.	✗			
2. Practises habits of cleanliness.	✗			
3. Has good postures.	✓			

Comments :

Teacher

Fig 9.2 : A sample report card

given in a report card. They do not mean the same thing to readers. These scores and grades should be supplemented by checklists as shown in Fig. 9.2.

In the elementary schools such diagnostic checklists are used to indicate the progress of a child in achieving the various objectives of teaching a subject. The following statements for reading and arithmetic illustrate the nature of these reports.

Reading :

1. Reads with understanding ______
2. Works out the meaning and the use of new words ______
3. Reads well to others ______
4. Reads independently for pleasure ______

Arithmetic :

1. Uses fundamental processes ______
2. Solves problems involving reasoning ______
3. Is accurate in work ______
4. Works at a satisfactory rate ______

The checklist form of reporting as described above, has the obvious advantage of providing a detailed analysis of the pupil's strengths and weaknesses, so that constructive action can be taken to help him improve his learning. The traditional report card does not provide any interpretation of strengths and weaknesses, does not show any relation of performance to potentiality. The checklist type of report form appears a commendable move toward providing the parents detailed information about the child.

The report card, howsoever modified it may be, does not convey all the information about the students to the parents with proper perspective. Hence it needs to be supplemented by other means such as informal letters and parent-teacher conferences. Both these latter methods have a greater flexibility in reporting. In the parent-teacher conference, an effective two-way communication is possible. The teacher can explain what the marks mean and why they were given. Parents can arrive at an understanding of their child's strong and weak points in each subject. They can be told whether he seems to be achieving in accordance with his ability, whether or not he is improving and what they might do to help him. The teacher in turn might learn from the parents significant facts about the child's home life and background which will help her understand and work with the child more effectively while he is in the school.

FOR FURTHER STUDY

1. Explain the terms reliability, validity, objectivity and adequacy of a test. How are they related to one another?
2. Discuss the various factors that affect reliability and validity of a test.
3. Why is it difficult to interpret test results by considering only the raw scores? What are the peculiar features of the yard-stick used in educational measurement?
4. What is the effect of the following factors on the reliability, validity and usability of a given test?
    (i) The test uses needlessly difficult language.
    (ii) The examiner has to use his opinion while scoring the test.
    (iii) The wording of the items is ambiguous.
    (iv) The test requires about two and half hours for complete administration to a class.
    (v) The examiner is allowed to give additional marks for good handwriting in a language test.
    (vi) The test is too short, consisting of ten items only.
    (vii) There are some questions which are of the nature of a quiz.
    (viii) The test uses a complicated procedure for scoring.
    (ix) Too much time is allowed in a speed test.
    (x) A written test is used to judge the laboratory skills of pupils.
5. Review the criteria of a good test and show why a good test must possess all of them if it is to serve its purpose efficiently.
6. Explain the terms: raw scores and derived scores. How do the derived scores help in interpreting test results?
7. State and explain the various types of norms.
8. What are the advantages and limitations of:
    (i) Age norms
    (ii) Percentile norms
    (iii) Grade norms
    (iv) Standard score norms.
9. What is the importance of assigning a mark or a grade in reporting pupil progress? What are the drawbacks in the traditional report card? How will you make it more meaningful?

10. "Raw scores by themselves are meaningless in reporting pupil achievement." Discuss this statement with examples and explain the different devices that are used to make the scores meaningful.

11. What do you understand by validity and reliability of a test ? How are they measured ? What is the relationship between the two ?

12. Which of the listed norms would likely be most suitable for use in each of the following tests ? Give reasons.
    (age norms, grade norms, percentiles, standard scores, and quotients.)
    (a) A test to be used for selecting the most intelligent students for admission to a college.
    (b) A test designed to indicate the degree of retardation or acceleration of the pupils in an elementary school, in arithmetic.
    (c) A test designed to furnish the level of intellectual development of a sixth grade child.
    (d) A test to indicate the relative mechanical aptitude of applicants for factory jobs.
    (e) A test designed to indicate the acceleration or retardation of a child's mental development.
    (f) A test yielding information about the acceleration or retardation of elementary school pupils in height or weight.
    (g) A test giving scores to indicate the relative honesty of a pupil.
    (h) A test designed to yield comparable scores of a pupil in arithmetic, spelling and reading.

13. A test constructor announces that he has constructed a highly valid test of intelligence. What information is needed to give this statement meaning ?

14. Distinguish between :
    (a) A norm and a standard
    (b) A raw score and a derived score
    (c) Reliability and validity of a test
    (d) Physical measurement and Educational measurement.

Answers :

Q. 4. Validity 1, 5, 7, 9, 10. Reliability 2, 3, 6. Usability 4, 8.
Q. 12. Age norms : c, e, f. Grade norms : b Percentiles : a, d, g.
Standard scores : h Quotients : c, e.

* * *

10

TOOLS OF EVALUATION : I

[Quantitative Techniques : Class-Room Tests]

We have discussed in the previous chapter the essential qualities of measuring instruments. For evaluation purposes we have to use a number of tools and techniques which will enable us to measure various learning outcomes such as :

1. Can Shekhar read well ?
2. Has he developed any interest in reading ?
3. Can he handle instruments in the laboratory ?
4. Does he co-operate with others ?
5. Does he use his time wisely ?
6. Can he solve problems in mathematics ?
7. Can he express effectively his ideas orally ?
8. Does he participate in game ?

For answering these questions we have to use a number of techniques. Some of these may be in the form tests while others may be of an informal nature which mainly rely on observing the pupils in various situations. These may be also termed as *quantitative* and *qualitative* techniques. The former type gives us numerical scores while the latter gives us verbal descriptions (Table 1.1, page 18). At present however main emphasis has been on the written examination. Public examinations are dominating the educational field in India. Frequently alarming reports about these examinations appear in newspapers. (Fig. 10.1).

Examinations in India have been severely criticised by all educators. The defects of the present system of examination have been repeatedly pointed out by various commissions and conferences from time to time. These defects have made the examinations the target of much public criticism, so much so that we often hear the voice of the extremist reformer, "*Abolish all examinations ?*" But this remedy appears to be more dangerous than the disease. Instead of abolishing the examinations, we may think of improving them.

WANTED
Supervisers for conducting Examinations
EDUCATIONAL QUALIFICATIONS – Nil
PHYSIQUE – Tall, strong & sturdy
SECURITY – Candidates will be insured for a minimum of Rs. 20,000 to secure them against any mishap.
FINAL EXAM.
DOON SCHOOL
S.S.C. EXAM.

"Raju, why do you need the gun?"
"Daddy, we have our examination today."

Fig. 10.1: Present day Examinations.

What are the defects in the present system of examination?: (i) Too much importance is given to acquisition of information while other important objectives of teaching are ignored. (ii) The questions set at the examinations are selected at random and are not comprehensive enough as a whole. (iii) Students interpret the questions differently, because of ambiguity of wording. (iv) The assessment of essay type questions by teachers suffers because of subjectivity and the judgment of different teachers about a student's performance becomes widely diverged. (v) Teachers think more of teaching a subject than the child. They teach whole essays by heart for reproduction in the examination. (vi) Undue weight is attached to the ability of a pupil to memorize relatively unimportant details learnt for the examination only. (vii) Instead of considering the examination as a means, it is considered as an end. Students and teachers lose sight of the real goals of teaching and learning a subject and all efforts are directed towards attaining high marks in the examination. (viii) Certain typical questions always repeat themselves. That is why there is always a tendency to guess the questions which are most likely to occur. Thus passing in an examination becomes a matter of chance. (ix) Pupil's progress is reported in terms of percentage of marks he scores in a subject which is just cold statistics and does not tell us how good his progress is. (x) Teachers are not able to identify the strengths and weaknesses of their own teaching from the way a pupil's performance is reported at present. (xi) Whatever is important for the examination is given importance in class-room teaching. The examination thus dictates *what to teach* and *how to teach* in the class-room.

The Net Result : The effect of all this, is a change in our values, concepts and attitudes !

1. Who is considered as a good teacher ? One who successfully leads his pupils through the examination and not the one who tries to inculcate good qualities among his students.
2. What is the function of a school ? It is a shopping centre where information is being sold to the customers. It is not looked upon as a centre of an individual's all round development.
3. Since the emphasis is on the supply of information, coaching classes are flourishing. In some places they have become

BOMBAY CHRONICLE
June 6, 1971
S.S.C. Examination
Results 28%
128,400 students fail.
Students say : 'Papers were difficult, unexpected questions were set'.
Demand for reassessment.

Join Our Lucky
Coaching Classes
Bombay 89.
Success Guaranteed !
Money will be refunded!
Concession in fees.
Contact : Phone 562814

Fig. 10.2 : Education to-day.

more popular than the schools. How many schools can face the competition of coaching classes if these classes are recognized by Government as on par with schools ?

4. Passing the examination has become the goal of education and all efforts are made in that direction only. Qualities such as co-operation, initiative, perseverance, tolerance etc. are neglected.
5. A student who passes the examination many times fails in the world of work for want of qualities mentioned above.
6. There is a tremendous wastage of human resources because of a large number of failures in public examinations (Fig. 10.2).

It will be seen from the above discussion that the examination is dominating the whole system of education and the various activities in any school. Naturally a critic may ask, "*Why not abolish the examinations ?*" It sounds well and some teachers may even vote in favour of this proposal. But as educationists we shall have to think twice before agreeing to such a drastic proposal. What purpose is served by the examinations ? What will happen if the examinations are abolished ?

Examinations have since long served as a great motivating factor in the life of an individual. Students will not study hard, if the external examination is removed. Teachers under training will assume *the take it easy attitude* if there were no examinations for them. It is also likely that teaching in the class-room will be adversely affected for want of any external control through the examination. Results of examinations supply useful information to all concerned with education. If examinations serve such varied purposes and give useful information to students, teachers, parents and headmasters, we cannot think of abolishing the examinations but we must think only in terms of improving them. We shall consider in this chapter various types of classroom tests, which give a quantitative measure of achievement.

The two main kinds of teacher made achievement tests are the essay or free-response tests and the objective type of tests. There is a difference of opinion as to the relative merits of each. Neither is perfect and neither is better than the other *in general.*

Each has its uses and limitations. The choice of either of the two types will primarily depend upon the objectives which are to be measured.

I. The Essay Type Examination

Essay questions should be used primarily for the measurement of those objectives which cannot be measured by objective test items. The main value of the essay examination is that it requires the student to organize his knowledge, to marshal his facts, arrange them logically and evaluate their relative importance. Abilities such as these, are not ordinarily tested in objective type of tests. The essay question demands that the students *generalize, compare, contrast, explain, criticize, summarize, analyze, illustrate, defend, describe* or in other words use the *higher mental processes.*

Merits : (i) The most important reason for retaining the essay type is that no substitute has been found that will measure the objectives such as *ability to organize ideas, ability to interpret, to criticize, to express* etc. (ii) This type encourages better study habits than objective tests do. Pupils develop such study habits as making outlines and summaries, understanding relationships among the facts, organizing the material etc. (iii) An advantage that is more apparent than real is the ease of construction of essay tests. They can also be easily administered. (iv) Another feature is the adaptability of the essay questions to school subjects. It is possible to use this type practically for all subjects. (v) The student is given liberty to respond in his own way to the question asked. The answers that are produced need not be completely right or completely wrong. All degrees of comprehensiveness and accuracy are possible.

Limitations : The limitations of the essay test are so severe that it would probably be discarded entirely as a measuring instrument were it not for the fact that it measures significant learning outcomes which cannot be measured by other means.

1. Let us first consider the factors that lower the *validity* of the essay test.

(i) The first major limitation of the essay test is its limited sampling of the content of the course. So few questions can be included in a given test that some areas are measured intensively while many others are neglected. This inadequate sampling is not representative of the course content hence it lowers the validity of the essay test. (ii) A second outstanding characteristic of the essay question is the subjectivity of scoring. A series of studies have shown that answers to essay questions are scored differently by different teachers and that even the same teachers score the answers differently at different times. Such results hardly foster confidence in the essay test. The subjectivity of scoring may be the result of varying standards of expectancy among the teachers concerned; one teacher may attach importance to the factual content, another may look more to the organization of ideas and the third to the style of expression. Thus different learning outcomes are being measured through the same answer; it is just like *three blind men* appraising the elephant. (iii) The essay examination is contaminated by such extraneous factors as spelling, handwriting, grammar and bluffing. The student with facility in verbal expression may write pages on a topic without saying anything. Every experienced teacher is familiar with the student *who says nothing* but *who says it very well*. In addition to skill in writing, quality of handwriting frequently influences the grading on an essay test. Students have been penalized because the examiner got irritated by poor handwriting. Spelling may be important in a language examination but penalizing a student for poor spelling in physics or in history may not be fair. The test would then measure *verbal ability* and not the *achievement in physics*. Similarly other factors such as grammar, vocabulary, neatness, organization of matter enter into the grading and lower the validity of the test. These factors may be given weight in grading but only when they are a part of the learning outcome and where pupils have been informed that the factors are to be given weight in grading. (iv) Another factor which affects the teacher in his marking is that he typically knows the pupils whose papers he is marking. He is probably prone to give pupils who have previously done good work or at least made favourable impressions on him, the advantage of the *halo effect*. This term is used when our judgment is biassed by previous impressions. (v) The essay questions are many times so worded that students do not know the exact implications of the question. They just guess and try to bluff, e.g., consider the question :

"Rewards and punishments are like a double edged weapon." Discuss.

It was found in one examination that many of the students did not understand the meaning and the implication of the statement; they wrote all about rewards and punishments. It was nothing but bluffing.

All the above factors introduce errors which lower the validity of the examination. A student having a good handwriting, consistently gets a higher score in English composition. If the high score is partly due to good handwriting, the score is not telling the whole truth about the achievement hence the validity is lowered.

2. Let us now consider other limitations which lower the *reliability* of the essay test. Many studies have shown that the answers to the essay questions are scored differently by different teachers. The factors that contribute to the subjectivity of scoring are rather obscure.

(i) The length of an answer is one that is often found to operate below the conscious level. Sometimes the teacher allows himself to be carried away by padding or unnecessary words included in the answer. It is a difficult fault to correct. (ii) The general feeling of well being or the opposite mood of the examiner is a normal phenomenon that influences grading. The student who said, "I hope my history professor does not quarrel with his wife and that he has a good breakfast before reading my paper" was reflecting some of the personal variables that affect scoring the essay test. (iii) Similar to the *halo effect*, it is equally possible to be the victim of what some authors have called the *horns effect*, that amounts to carrying an impression of an inferior answer over to all the rest of the work a student has done, some of which may be fairly good in quality. (iv) Comparison of one paper with others may result in an improper grading; a mediocre paper looks very good if the teacher has just read a series of poor papers while the same answer seems very poor if it is preceded by a series of excellent papers. (v) It is difficult to evaluate the answers objectively simply because the answers are of all degrees of correctness. The responses vary in many ways and by infinitely small degrees. Hence there is a wide disparity in scoring. (vi) Sometimes the wording of a question is ambiguous. It is interpreted differently by various students. The same student may interpret it in another way on a second occasion.

All the factors mentioned above affect the scoring of the test which in turn lowers the reliability of the test.

3. *Low Usability* : This type of test is time consuming. When the class is large, the task of reading the papers is an extremely laborious one and conscientious scoring becomes practically impossible. If we want to improve the grading it will require even more time for assessing the essay examination. The expenditure of time and energy by both pupils and teachers is such a serious limitation that essay examination should be used only for those objectives that cannot be measured with objective tests.

Improving the Essay Examinations

Essay examinations will continue to be used in evaluation of pupil performance because of the ease with which they can be prepared and because of their advantages. We cannot think of discarding the essay type questions but should think in terms of improvement. We can improve this examination in two ways : (1) improving the construction of essay questions and (2) improving the scoring.

1. Suggestions for Improving Essay Questions

(i) Restrict the use of essay questions to those learning outcomes which cannot be satisfactorily measured by objective items. Such functions as ability to organize, to express, to interpret and to elicit understanding may be tested through the essay questions. **(ii)** Do not start the essay questions with such words as *what, who, when, enumerate*, etc. These words are likely to present tasks requiring only the reproduction of ideas. In general, start the essay questions with such phrases as *compare, contrast, discuss, explain* etc. Consider the following questions :

1. What are Thorndike's laws of learning ?
2. What is the effect of overlearning on retention ?
3. When do we find a plateau in learning ?
4. Enumerate the causes of juvenile delinquency.

In all these questions, only the *knowledge objective* has been tested. Students have to reproduce the information learnt by heart. We may modify two of these questions so that other important objectives can be assessed. We may retain the other two for measuring the knowledge objective.

1. Explain how Thorndike's laws of learning can be applied to class-room situation.
2. How will you explain the fact that we remember the nursery rhymes even after a period of 10 or 15 years ?

(iii) Write the essay question in such a way that the task is clearly and unambiguously defined for each examinee. The question should not be too vague, general and comprehensive. Consider the following questions :

1. *Poor* : What is the bearing of psychology on education ?

This question is so general and comprehensive that different students will emphasize different aspects while the examiner may have altogether a different view. The question may be broken up into shorter but specific questions or the scope of the answer may be defined, as shown below :

Better : Discuss the bearing of psychology on education with reference to (a) methods of teaching and (b) the curriculum.

2. *Poor* : Compare the essay examination with the objective type of examination.

Better : Compare the essay examination with the objective examination with respect to (a) learning outcomes measured, (b) framing of questions, (c) scoring and (d) influence of extraneous factors.

In the modified form, the task in both the above questions is more clearly and unequivocally defined for each examinee. Questions should not be in the form of a quiz. Otherwise students try to guess the meaning and write the answer.

Poor : "Instincts are the raw material of character and throughout his task the educator must deal with them." Discuss.

Better : Why is it said that while dealing with children, the teacher should make an appeal through the instincts ?

(iv) Directions for the test should be explicitly written. Here is an example of a certain language examination.

(a) Each question carries 20 marks.
(b) Marks will be deducted for writing beside the point.
(c) While assessing the answers, importance will be given to the candidate's ability to organize the matter and to critically examine the various view points.
(d) Marks will be deducted for mistakes in spelling and grammar.
(e) Each answer should be restricted to about 600 words.

(v) Avoid the use of optional questions. A fairly common practice in the use of essay questions is to provide pupils with more questions than they are expected to answer. When pupils answer different questions, it is obvious they are taking different tests and the common basis for evaluating their achievement is lost. **(vi)** Students are found to be misinformed about the meaning of important terms used in essay questions. Students frequently *discuss* or *describe* when asked to *define;* they may also *elaborate* in great detail when asked to *summarize.* A solution would be to supply the necessary training to the students in writing essay questions. **(vii)** Allow liberal time limit so that the essay test will not be a test of speed of writing. While setting up the question paper see that it can be answered in the allotted time, leaving some time for reading the questions, drawing up an outline of the answers and finally for revision. **(viii)** We have seen that the essay examination suffers from lack of adequate sampling. This defect can to some extent be overcome by increasing the number of questions and by limiting the length of their answers. A question paper with 10 questions would represent a better sample than one with 5 questions only. **(ix)** One of the favourite questions of examiners is :

Write short notes on

One can ask anything under this heading. The following topics have appeared in various question papers.

Poor : *Write short notes on* :

(a) Social Education	(b) Home Work
(c) Heuristic Method	(d) Teaching Aids.

This type of question is worse than the ones mentioned in (iii) above. The student does not know the limits and he goes on writing pages after pages. You can define the scope by using such wording as :

Better : *Write short notes on* :

(i) Objectives of social education.
(ii) Misuse of home work.
(iii) Advantages of heuristic method.
(iv) Effective use of teaching aids.

2. Suggestions for Scoring the Essay Examination

(i) Prepare an outline of the expected answer in advance, showing what points are required and the credits to be allowed for each. This will provide a common frame of reference for evaluating the individual papers. (ii) Decide in advance what factors are to be measured. If the ability to organize, to interpret or to apply the principles is to be assessed, the examiner should not allow himself to be biassed by bad handwriting, spelling, sentence structure or neatness. The ability to write, to spell or to use correct English can be assessed through other suitable tests. (iii) Examination should be scored as far as possible by the one who frames the questions. He is the person who can have a clear picture of the expected responses. Whenever more than one examiner are involved, they should be brought together to develop a uniform scoring procedure. Model answers and marking schemes may be discussed and finalised in this meeting. (iv) Grade the papers as nearly anonymously as possible; the less you know about *who wrote an answer*, the more objectively you can grade *what was written*. (v) Score one question through all of the papers, before considering another question. This type of scoring permits the examiner to concentrate on the answer to a single question and to judge better the merits of the several pupil responses to the same question. (vi) When important decisions such as selection for awards or scholarships are to be based on the results, obtain two or more independent ratings and average them.

When these instructions are followed, *subjectivity* in scoring will be reduced and we will be nearer the goal of *objectivity* in scoring.

II. The Objective or New Type Tests

Objective or short-answer examinations have increased in popular usage over the years. Factors that have contributed to the increased frequency of their use are the low reliability of essay tests and the large classes which increase the load of reading a large numbers of essay papers.

An objective test is highly structured and requires the pupil to supply a word or two or to select the correct answer from among a limited number of alternatives. The word 'objective' in *Objective Tests* means that the scoring will not vary from examiner to examiner. An objective test may not be necessarily *objective-centred* though it can be *objectively* scored; an essay test however, may be *objective-centred,* though it may be difficult to score it *objectively.* When the questions in an essay or new type examination are framed with reference to the objectives of instruction, the test becomes objective-centred, but when the questions are framed aimlessly, the test may not be objective-centred. A few examples of objective-based questions from various school subjects are given in Appendix I and II.

It should be clear then that tests of any kind should be based on the *objectives* of instruction, also that tests of any kind should be made as *objective* as possible.

Merits of Objective Tests

(i) *Extensive Sampling* : One of the criticisms of the essay examination is its limited sampling. But in the new type test, since each item is brief, many items can be included in the test. These items can be spread evenly over the topics to be covered and a more representative sampling can be obtained. This reduces the role of luck; candidates cannot spot out and cram the likely or expected questions, as is done in the essay type. All the hundred or more questions of the new type can probe the strong and weak points of the students.

(ii) *Objectivity in Scoring* : In an objective test, the items are so stated that the answers are brief and definite; in a well prepared test only one correct answer is possible; such a test can be scored repeatedly by one person or it may be scored by a large number of different persons with practically no disagreement in the scores assigned. With a predetermined key, the score will be the same whosoever scores the test; the score will be the same, no matter who it was that wrote the answers.

(iii) *Economy of Time* : The form of the objective items makes it possible for the pupil to record his answer definitely and briefly. Hence a wider area of content can be covered in a given period, unlike the essay examination. Moreover, the scoring of the tests can be done very accurately and speedily. Scoring can be done even by electrical test scoring machines.

(iv) *Elimination of Irrelevant Factors* : Factors such as speed of writing, literary style, fluency of expression etc., distort the examiner's marking in the essay examination, but such is not the case in the new type test. The fast writer does not get unfair advantage over the slow one, because writing by the pupils is reduced to a minimum. A further advantage is that the examinee who often gets several marks by bluffing, is entirely foiled by the new type tests.

Limitations of Objective Tests

(i) *Measuring Trivial Aspects of Ability* : It is said that this type of test inadequately allows opportunity for the pupil to organize and express his thoughts. The test measures only the most trivial aspects of ability such as details of information and fragments of knowledge. Such tests cannot, as can the essay type of examination, show the examinee's general understanding of the subject, nor his interpretation of facts, his capacity for organizing and formulating his knowledge, nor his initiative and originality. The criticism of triviality and fragmentariness seems at first sight justifiable and serious because many new type tests are incompetently constructed. Many teacher-made tests do not contain thought-provoking items; but that does not mean that they cannot be made to do so, when teachers become masters of objective techniques. There is nothing inherent in new type examinations, which makes better questions unattainable. Of course examinees have more opportunity of displaying higher mental processes in essay-type than in most new-type examinations.

(ii) *The Guessing Factor* : Examinees may obtain a certain number of correct answers by random guessing. This becomes a problem particularly for true-false questions in which there are only two choices. The problem of guessing would be serious in a short test with few answer choices, but chance success will be reduced by including a large number of items. All students will adopt the same policy, if instructions about guessing are made sufficiently definite.

(iii) *Cost* : Cost of printing a new type test is considerably greater than that of an essay type examination. But this cost represents a very minor item of expense when considered in terms of the real educational importance.

(iv) *Difficulty of Preparation* : Scoring the new type test is a time saver for the teacher, but this advantage is offset by the time required to construct adequate objective questions. The formulation of questions that meet the requirements of being completely true or false or that have a best answer is very difficult. But considering the time saved in scoring and the reliability of results obtained thereafter, it is worthwhile spending time in the construction of the test.

Both the essay and objective tests can be used to evaluate pupil achievement. We must remember that both types are imperfect mental measuring instruments. By combining the two, their draw-backs partially cancel one another and the total validity is definitely superior to that of either of the separate types. We shall discuss the various types of objective tests in the next chapter.

III. Oral Examination

It is the oldest form of examination used by teachers. In old days teachers relied wholly on this form. The modern teacher rarely attempts to grade every oral contribution of her pupils. When we think of the oral examination, the typical teacher conjures up a picture of a classroom situation in which the teacher spends a class period asking each pupil one question in turn and then assigning a grade for each student. Obviously such a practice has little measurement value.

Merits of Oral Examination : (i) The oral questioning of an individual pupil constitutes an excellent means of following the thought processes which he has used in solving certain problems, such as one in mathematics. Used in this manner, it is a valuable tool for the diagnosis of pupil difficulties. (ii) This type can be used in determining how well an individual pupil has integrated his knowledge, how well he can apply it to various situations and can see its implications. (iii) Certain types of performances such as oral expression, pronunciation, and performance in debates can be evaluated only through oral examination. (iv) Oral examination constitutes the only way of measuring achievements in the case of young children who have not yet mastered the intricacies of reading and writing. (v) The relative informality of the oral test situation is far less frightening to the timid child.

Limitations : (i) The oral examination is not equally fair and just to all pupils. (ii) It does not test extensively or efficiently. (iii) It permits interference and favouritism by the teacher. (iv) It is unjustifiably time consuming. (v) It leaves no permanent objective record of pupil performance. (vi) It does not permit an evaluation of the difficulty of questions. Item analysis is difficult to obtain. (vii) Sampling is inadequate as only 1 or 2 questions are asked; this lowers the reliability of the oral examination.

Suggestions for Improving the Oral Examination : (i) Much of the misuse of the oral examination is probably due to poor planning. Decide the objectives and the content areas for which the oral examination is to be used. (ii) This type should not be used for measuring knowledge of specific facts which can be readily measured by a written objective examination. (iii) Write down the questions that you want to use. (iv) Acceptable answers should be written for each of the questions constructed. (v) The method of grading should be decided upon before the test is given. (vi) When only one question is asked to each of the pupils, care must be taken to keep all questions comparable in difficulty.

If these precautions are taken, much of the subjectivity in grading will be reduced.

IV. Performance Tests

Tests of the types thus far discussed are often inadequate to measure some areas of achievement. For example, a pupil's achievement is sometimes expressed by means of a *product* or a *procedure*, which is a direct indication of the pupil's skill and understanding. Performance of any task includes two important aspects : (i) the procedure, skill or technique and (ii) the product or result. Performance testing is valuable for any subject in which pupils are being taught to follow specific procedures or to create some product. Actually this type of testing is useful in a wide range of subjects such as English composition, maps, drawings, models, in which case we get the *products*, while *procedures* are emphasized in the field of speech, music, physical education, laboratory experiments, and craft work.

Tests of performance may be classified as (i) object tests, (ii) performance measures and (iii) product evaluations. Performance tests are also used for measuring intelligence and aptitudes. The tests discussed in this section are used as evidence of educational achievement.

(a) *Object Tests* : They measure knowledge, concepts and understandings which are prerequisite to certain skills. They are sometimes called identification tests because the pupil is asked to identify specimen etc. The specimen can be presented in actuality or in the form of photographs.

(b) *Performance Measures* : The process by which a pupil produces some type of desired result is appropriately observed and evaluated as to quality by use of checklists and as to quantity by noting the time required, for example, use of a microscope by a student, making proper connections in a circuit, preparing a piece of woodwork or metal work etc. In these situations the procedure will be mainly observed.

(c) *Product Evaluation* : The quality of a skill performance can be evaluated in terms of the characteristics of the completed product. In assessing the product, quality scales, rating scales and quantitative techniques are usually employed. In typing for example, the quality of the typed material is judged by counting the number of errors as well as number of words typed per minute. Handwriting and pieces of fine arts or woodwork are evaluated by using quality scales as well as rating scales.

Table 10.1 : Strengths and Weaknesses of various test forms

Strengths	Weaknesses
Oral	
(1) Most valuable tool in evaluating oral expression and pronunciation	(1) Time consuming
(2) Useful as a diagnostic tool, useful with young children who fear written examination	(2) Poor sampling of a pupil's knowledge
(3) Useful in day-to-day teaching	(3) Not equally fair and just to all pupils
	(4) Is not subject to refinement through item analysis
Objective	
(1) Gives an extensive test sample	(1) Frequently neglects measurement of higher thought processes
(2) Can be made highly reliable	(2) Overemphasizes rote learning
(3) Can be graded objectively and quickly	(3) Promotes poor study habits
(4) Eliminates bluffing and avoids the influence of extraneous factors such as style, and legibility	(4) Is difficult to prepare
(5) Can be refined through item analysis	(5) Costs more than essay test to prepare and to reproduce
(6) Can be adapted to several teaching objectives	(6) Encourages guessing
(7) Can be made highly valid for some teaching objectives	
Essay /Teacher made	
(1) Useful for measuring understandings, thinking and other complex learning outcomes such as organization of ideas, and originality	(1) Gives a limited test sample
(2) Relatively easy to construct	(2) Is difficult to grade
(3) Encourages proper study habits	(3) Influence of extraneous factors such as style of writing, spelling, handwriting, etc.
(4) Adaptable to several school subjects	(4) Encourages bluffing
	(5) Has low reliability
	(6) Influence of halo-effect
Performance	
(1) Stresses application of knowledge	(1) Is not adaptable to many fields of learning
(2) Can be used as a learning device	(2) Is difficult to construct
(3) May give a true picture of a verbally handicapped pupil	(3) Is often difficult to grade
(4) Measures some skills, not measured by other test forms	(4) Is often time consuming

V Standardized Tests of Achievement

The informal class-room tests discussed so far are useful in measuring the achievement of pupils, but these tests have their limitations also. A teacher, for example, cannot easily compare the performance of her pupils with that of other pupils from neighbouring schools with the help of the routine class-room tests. A standardized test can help her do the comparison. Such a test is based on the general content and objectives common to many schools spread over a wider region and the test publisher reports the scores, these larger groups have made on the test. Thus a given score can be compared not only with the scores in one class or one school, but with the scores of the larger group. A teacher-made test reflects the content and the objectives specific to her own class, while a standardized test has a broader coverage and measures outcomes and content common to a large number of schools

Achievement Test Batteries : Standardized achievement tests are frequently used in the form of test batteries. A battery consists of a series of individual tests all standardized on the same representative group of pupils. This makes it possible to compare test scores on the separate tests and thus determine the relative strengths and weaknesses of pupils in different areas covered by the test. With an elementary school battery, for example, it is possible to determine that a pupil is strong in language skills but weak in arithmetic skills, good in reading but less proficient in spelling and the like. Such comparisons are not possible with separate tests which have been standardized on different groups of pupils, since the base for comparison is not uniform. When an achievement test undertakes to provide a general overall appraisal of status in some subject it is called a *survey test*. It usually gives one score which shows the pupil's status in that particular subject.

VI Diagnostic Tests

A diagnostic test unlike the survey test, undertakes to provide a detailed picture of strengths and weaknesses in an area. The survey test is not useful for locating specific weaknesses of the pupils, but it serves as the initial step to identify pupils who need diagnosis. A survey reading test tells us that Johnny who is starting the fourth grade, performs on our test at a level typical of a child beginning the second grade. This only provides a general appraisal of status in reading but a diagnostic test may indicate his specific weaknesses, for example, Johnny has a fairly good vocabulary, but he has no skill for working out unfamiliar words, that he makes frequent reversal errors, he has an above-average speed in reading but he fails to note details while reading. These findings together with others provide the basis for planning remedial teaching of reading.

Education borrowed the term *diagnosis* from medicine, where its fundamental character has long been recognized. The physician does not cure in groups, people who may be apparently having the same illness. He uses a stethoscope, X-rays, and blood tests to obtain evidence for making better diagnosis than he could otherwise make. These aids to diagnosis are therefore known as *diagnostic instruments* or *tests*. The teacher, like the physician, uses tests as instruments to obtain evidence for making better diagnosis of pupil difficulties.

The above discussion will make it clear that the development of diagnostic tests involves two steps : (i) analysis of the complex performance into its component subskills and (ii) developing tests for component skills.

Let us consider the fundamental operation of *addition* in arithmetic. It is not a single unitary ability but a composite of many elements. This process if analysed, may indicate elements of the following type :

> Simple combinations, higher decades, carrying numbers, adding columns, adding numerators with common denominators, adding mixed numbers, adding fractions, and decimals etc.

When the test is designed to measure all these elements it is possible to detect the errors, committed by pupils. In day-to-day teaching, if a teacher tries to observe the errors of pupils, she can make diagnosis of pupils' difficulties. Following examples will show the nature of mistakes, committed by pupils.

$$\begin{array}{r} 2\frac{1}{6} \\ +\,3\frac{1}{3} \\ \hline = \end{array}$$

In a class of 40 pupils, 28 missed the following problem of addition. On analysis it was found that 20 pupils wrote the answer as $5\frac{3}{6}$; the mistake was the failure to reduce the fraction to its lowest terms; four pupils gave the solution as $5\frac{2}{9}$: it is apparent that these four students merely added the numerators and denominators without reducing the fractions to a common denominator. The remaining four pupils got the answer as $5\frac{2}{6}$, which represents two types of errors. This author has come across some errors of a peculiar nature in various fundamental processes in arithmetic as described below:

(1)
$$\begin{array}{r} 636 \\ -\,363 \\ \hline 333 \end{array}$$

(2)
$$\begin{array}{r} 37 \\ 83 \\ 94 \\ 45 \\ \hline 43 \end{array}$$

(3) (a) $\frac{3}{4} \times \frac{4}{5} = \frac{3}{5}$

(b) $1\frac{1}{2} \times \frac{2}{5} = 1\frac{1}{5}$

(4)
$$\begin{array}{l} 6)3647(67 \\ \quad \underline{36} \\ \quad\;\; 47 \\ \quad\;\; \underline{42} \\ \quad\;\;\; 5 \end{array}$$

(5)
$$\begin{array}{cccc} (a) & (b) & (c) & (d) \\ 38 & 72 & 25 & 56 \\ \times 7 & \times 8 & \times 6 & \times 7 \\ \hline 566 & 646 & 300 & 632 \end{array}$$

In the first example, the child has a vague idea that the smaller digit is to be always substracted from the bigger one. In the second example which is on addition, the child merely added all the digits, irrespective of their place value. In the third example, the child carries over the skill used in 3 (a) to the example 3 (b), without considering the nature of the fractions given. In the fourth example, the child does not understand the role of zero as a place holder while determining the quotient. In the fifth example, the girl was not doing the process of carrying the digits properly. She added the number to be carried to the *tens* digit in the multiplicand before multiplying.

In all the above cases it will be seen that improvement will not take place unless specific weak spots are located. Two students may be found to be weak in the fundamental operation of addition, it does not mean that both of them have the same difficulties. Remedial measures as commonly used by teachers many times fail for want of proper diagnosis. Students mentioned above will show definite improvement when their specific weak points are attended to.

In the teaching of English, we hear the remark that pupils are weak in *spelling* the words. We can minimise the spelling errors, if we try to analyse and classify the errors. Following broader classification is an attempt of diagnosis.

(i) *Writing double letters for a single letter* : controll, cheerfull, orrange, skillful, writting.

(ii) *Single letter in place of double letters* : alright, runing, writen. puting.

(iii) *Misplacement of letters* : childern, alphebat, frist, lerant..

(iv) *Omission of letters* : blackbord, begining, oraly, reasnable, intresting, diffrent, envirnment.

(v) *Confusion and substitutions* : backword (backward), there (their), dose (does), vovel (vowel), storeys (stories).

(vi) *Insertion of letters* : boaring (boring), similiar (similar), wronge (wrong), diagrame (diagram), wheather (whether), heighest (highest).

The two areas in which diagnostic tests are most common are reading and arithmetic. The tests provide some rough and quite tentative hypothesis as to the individual's strengths and weaknesses. Remedial measures should be used on the basis of the results of diagnostic tests.

Uses of Achievement Tests

1. *Grouping of Pupils* : When a school is large enough to have several sections at the same grade level, a decision must be reached on some grounds regarding the classification. In some secondary schools grouping is done on the basis of the level of achievement.

2, *Group Comparisons* : Teachers generally wish to know just how their pupils achieve, in comparison with other similar pupils and classes. Standardized tests provide norms which enable teachers to make comparisons.

3. *Evaluation of Teachers* : In some schools standardized achievement tests are used either openly or covertly to evaluate the success of the teacher. This practice is seldom

advisable. It fails to take account of a number of factors that affect students' achievement. The achievement depends upon the ability of students, out of school experiences, methods of instruction, motivation and a number of such factors. Any one of these factors may have a more potent influence upon the achievement than the teaching ability of the teacher. If this practice is followed, it will have harmful effects on the attitude of the teacher.

4. *Selection* : While admitting new students, the Head Master can use these tests as a screening device.

5. *Grouping Children within a Class* : One of the effective techniques teachers have developed for dealing with individual differences in pupils is to form within the class, small groups of pupils who have about the same level of skill. Pupils in these groups may work on the same materials and at the same speed. Achievement tests can be used to form such groups.

6. *Diagnosis* : A teacher may wish to diagnose a pupil's weaknesses and gaps in the knowledge, understandings and skills. This will help her to pinpoint individual remedial work and to revise the methods of teaching with proper emphasis on weak areas. Remedial work will be successful if accurate diagnosis is done.

7. *Locating underachieving Students* : A teacher has to evaluate the achievement of the pupils in relation to their potentiality as shown by an intelligence test. She comes across students who appear to be underachieving in spite of their abilities to achieve more. She can make use of both these types of tests to locate such cases.

8. *Counselling* : Results of achievement tests are of great value to the counsellor in helping a student decide the future career. A student's performance on a mathematics test, for example, certainly has a bearing on the possibility of his succeeding in a science or an engineering course.

9. *Motivation* : Examinations in general motivate the students to do better work. Students accept success in an examination

as a worthy intermediate goal toward which they might be willing to work.

10. *Evaluation of the Curriculum and the Methods of Teaching* : By comparing the test performance of students over a period of some years, we can evaluate the curriculum and suggest changes if any. Results may indicate that the syllabus is too easy or too difficult or in some other way unsuited to the students. The results of achievement tests can be used to compare the effectiveness of different instructional methods.

FOR FURTHER STUDY

1. Compare the essay type examination with the objective examination with respect to (i) scoring (ii) framing of questions and (iii) influence of irrelevant factors.
2. Discuss the merits and limitations of essay examination. Suggest five guiding principles to improve this examination.
3. Criticize the following plan for an essay test in Geography :
   - (i) Time for the test is 40 minutes.
   - (ii) There are 8 questions, out of which 5 are to be attempted. Each question carries 16 marks.
   - (iii) 20 marks are reserved for neatness and good handwriting.
   - (iv) Marks will be deducted for grammatical mistakes.
   - (v) There will be 4 examiners to assess the answer-books.
4. Discuss the importance of oral and performance tests in a programme of evaluation. What are the limitations of these two types of tests ?
5. What is a diagnostic test ? Give examples to show how a teacher can make diagnosis of pupils' difficulties.
6. Discuss the uses of achievement tests in a school situation.
7. Criticize and revise each of the following essay questions.
   - (a) "Instincts are the bricks out of which the individual's character is formed." Discuss.
   - (b) Basic education is the application of the principle– 'education through activity.' Discuss.
   - (c) Write short notes on :

     (i) Intelligence and occupation, (ii) Education for citizenship, (iii) Education as preparation for life and (iv) Education as an art.

8. Indicate the type of test to which each of the following statements refers, using the symbols :

E for essay test O for objective test

(i) Best measure of ability to organize and express ideas about a subject ...... ( )

(ii) Best agreement among several persons grading the test ...... ( )

(iii) More susceptible to extraneous clues to correct answer ...... ( )

(iv) More chance for guessing ...... ( )

(v) Takes more time to construct ...... ( )

(vi) Students may not know how much detail to include ...... ( )

(vii) Gives better sampling of material on which the test is based ...... ( )

(viii) The student who knows the views of the teacher gets a higher score than the one who does not know ...... ( )

(ix) Advantage of the 'halo-effect' is possible ...... ( )

(x) Irrelevant factors such as handwriting and neatness do not affect the marking ...... ( )

* * *

11

TOOLS OF EVALUATION : II

(Quantitative Techniques : Objective Type Tests)

Writing Items for an Objective Test

The construction of good test items is an art. There is a misconception that objective test is the only remedy for improving the examination and teachers include a few objective tests in their question papers. After a careful analysis of these papers, this author has noticed that many of these objective type questions are prepared in a slip-shod manner without giving much thought to their effectiveness. Many of these test items simply measure the knowledge aspect. Such poorly constructed tests do not serve any purpose.

Scoring the objective tests poses no problem. Machines can do the job. A monkey can be trained to do it. The difficulty with such tests is in their construction. The principal types of objective test items may be listed as follows :

(i) Short-Answer, (ii) Completion, (iii) True-False,
(iv) Matching and (v) Multiple-Choice

These can be classified into two groups : The *recall type* and the *recognition type*. In the former the desired answer consists of a single word or a few words which the examiner must write in the space provided while in the latter, the examinee has to identify the right answer from among two or more answers printed on the examination paper. We shall first go over some suggestions that apply to almost any type of objective items. Then we shall consider the maxims for writing down the various types of test items.

General Suggestions for writing Objective Items

1. *Difficult words should be avoided* : Care should be taken to see that the words used in objective items are known

to all pupils so that no one will be handicapped because of a difficult word in the item, unless of course the purpose of the test is to measure *verbal ability*. Do not use metaphors.

Poor : What will be the cost of a score of saucers at six rupees a dozen ?

Better : What will be the cost of 20 dishes at six rupees a dozen ?

Students from the lower classes would be handicapped because of the two difficult words in the first statement.

2. *An item should not be reproduced verbatim from a text-book* : Sometimes a pupil has a memory for specifics of what he has read or heard which would enable him to answer the item from memory rather than in terms of *knowledge* and *understanding.*

3. *Authority must be mentioned for statements that might be considered controversial* : Ordinarily statements of controversial nature do not make good items, but there are instances where knowing what some particular person or school thinks, may be important for its own sake.

4. *An item should not provide cues to the answer of another item* : The first item below provides a cue to the second.

Example : The intelligence test considered most appropriate for individual testing of adults is

(i) The Stanford Binet, (ii) Performance Test by Bhatia.
(iii) Otis Group Test, (iv) The Wechsler Bellevue Scale.

Now consider a true-false item from the same test; students will easily know its answer, if they read the above item.

(*T.F.*) No test has been devised so far, to measure adult intelligence individually.

5. *Answers should follow a random pattern* : If the examinee discovers any pattern in the sequence of answers, he will guess at the answers with fair accuracy. If items are arranged alternately *true* and *false* or two true and two false, for example, the pupil is likely to discover the arrangement. Hence random occurrence of correct answers should be employed.

6. *Avoid ambiguous statements* : Care should be taken to see that each test item is subject to one and only one interpretation.

Poor : The shortest day of the year is in
(a) March, (b) June, (c) September, (d) December.

Better : The shortest day of the year in the northern hemisphere is in
(a) March, (b) June, (c) September, (d) December.

7. *Quantitative rather than qualitative words should be used* : Quantitative language conveys more exactly the meaning intended. Expressions such as *few, many, large, short* etc., are vague and indefinite. Mark T or F in the following statements :

Poor : (T.F.) There is little variation in IQ when an intelligence test is administered to the same individual on two occasions.

Better : (T.F.) Variation in IQ up to 5 points may be expected when an intelligence test is administered to the same individual on two occasions.

The answer to the latter statement is true. But the answer to the first depends upon how you interpret the term *little variation.*

8. *Avoid using interdependent items* : The answer to one item should not be required as a condition for solving the next item. Otherwise such dependence places more than the intended amount of weight on the first item. Every individual should have a fair chance at each item as it comes.

Example : (1) If a test shows consistency in measurement, it is called a————test, and if it shows truthfulness, it is called a————test (Ans. reliable, valid).

(2) The first of the above two characteristics of a test is determined by the————method. (Ans. test-retest).

In the above example the person who does not know the answer to the first question is at a disadvantage while answering the second item.

9. *Avoid trick and catch questions* : There is no justification for the inclusion of catch words, misleading statements or irrelevant confusions. Minor changes in spelling of wellknown proper names or the substitution of B.C. for A.D. after a historical date are an illustration of the tricky or catch questions. Such errors might be treated as *typographical* errors by some pupils; the rapid reader is likely to be penalized. Such items would be acceptable, however, if the test is designed to measure accuracy of proof reading.

Example : Water boils at 212°F and freezes at 32°C.

Here the student is likely to miss the inconspicuous letter C at the end of the statement, and thus is likely to be tripped. It is better not to have such an item.

10. *Items dealing with trivial detail should not be used* : An item should appraise some important item of knowledge or some significant understanding.

Example : The first psychological laboratory was established in the year (1) 1856, (2) 1879, (3) 1901, (4) 1916.

Undue weight is attached to the ability of the pupils to memorize relatively unimportant details, if such items are included in the test. As far as possible such an item should not be used.

We shall now consider the different types of objective test items.

I. Short-Answer and Completion Items

The short-answer item and the completion item are both of *recall* type, which can be answered by a word, a phrase, a number or a symbol. They are essentially the same, differing only in the method of presenting the problem. In the case of the short-answer item, a direct question is used while the completion item consists of an incomplete statement.

Example

Short-Answer : What is the name of the man who invented the steam engine ?

Completion : The name of the man who invented the steam engine is————————

Uses of these two types of items : Items of this type are well suited to testing knowledge of vocabulary, names, dates, identification of concepts and ability to solve numerical problems. Relatively simple learning outcomes are measured.

(i) *Knowledge of facts and principles* :
A word that qualifies a noun is called a(n)————————

(ii) *Ability to solve numerical problems* :
How many mangoes can I buy for Rs. 20 when they cost 40 paise each ?

(iii) *Simple interpretation of data* :
In the number 842 what value does the 8 represent ?

Merits : (i) Both these types are easy to construct; this is probably due to the relatively simple learning outcomes measured by them. (ii) A more important advantage arises from the fact that the pupil must supply the answer. This reduces the possibility of guessing by the student. He must either recall the information or make the necessary computations to solve the numerical problem presented. (iii) They are useful to measure the ability to interpret diagrams, charts and graphs.

Limitations : (i) These two types are best adapted to the measurement of rather highly factural knowledge of the type *who, what, when, where.* Both are unsuitable for measuring complex learning outcomes such as ability to apply, to perceive, to detect errors in the procedure etc. (ii) As compared with the other types of objective items, the short answer type is harder to score as it is not entirely objective. For example, a question such as :

Where was Jawaharlal Nehru born ?

The answer could result in the name of the city, district, state or country. Although the teacher had the name of the city in mind when he wrote the question, he cannot dismiss the other answers as incorrect.

Maxims for Constructing Short Answer and Completion Items

1. *Response should be definite* : The wording of the statement or of the problem should be so framed that the response will be a word, a date, a number or a short phrase. This will ensure objectivity in scoring.

Poor : What kind of process is evaporation ?
Better : To what does a liquid change when it evaporates ?

2. *The question or the incomplete statement should be so worded that there is only one correct answer* : Sometimes this is difficult to achieve. In such a case, all acceptable answers should receive credit.

Example : What method is used to preserve meats which are transported over long distances ?

Here the correct answer may be *refrigeration, cooling, canning* or *cold-packing.*

3. *Avoid over-mutilation of a statement* : Do not insert too many blanks. See that sufficient words are left to make it an intelligible question. For the same reason do not put blanks near the beginning of a sentence. Overmutilation reduces the task of the examinee to a guessing game or an intelligence test.

Poor : The ——— is obtained by dividing the ——— by the ———

Better : The IQ is obtained by dividing ——— by ———

In the first statement one can use the words *quotient, dividend* and *divisor* or *IQ, MA* and *CA*.

4. *Indefinite article before a blank should be written in the form a(n) so that the pupil does not get a clue* :

Poor : A piece of land entirely surrounded by water is called an ——— [Answer : Island]

Better : A piece of land entirely surrounded by water is called a(n) ———

In the first example an irrelevant clue is provided by the article 'an'.

5. *Make the blanks of uniform length* : If the blanks vary in length, they offer a clue to the length of the correct answer. Blanks should almost always call for but one word. If there are two consecutive missing words, it is better to have two blanks.

Poor : Air is ———, tasteless and odourless; without ——— we would die. Air forms the ——— the earth.

Better : Air is ———, tasteless and odourless; without ———, we would die. Air forms the ——— the earth.

6. *Blanks should not be used to fill up trivial details* : Only important and crucial aspects of a statement should be omitted.

Poor : The items in a(n) ——— should be arranged in an ascending order of difficulty.

Better : The items in a test should be arranged in a(n) ——— order of difficulty.

7. *Avoid lifting statements directly from the text* : It puts too much emphasis on rote memory and parrot like learning.

Poor : Much learning is by trial and ———.

8. *If the problem requires a numerical answer, indicate the units in which it is to be expected* : This will simplify the problem to the pupil and will simplify the scoring.

9. *A direct question is generally more desirable than an incomplete statement* : This is especially important for younger pupils.

Poor : America was discovered in the year ———.
Better : In what year was America discovered ?

10. *For purposes of easy marking it is desirable to arrange for answers to be written in the right hand margin* : It is difficult for the eye to pick out answers scattered all over the page. The scoring is more rapid if the blanks are numbered and students are directed to write the responses in the correspondingly numbered blank in the answer column at the right.

Example : Sound waves travel through 1—— at about 1100 feet per second. The speed 2—— when the temperature rises. Hence you would hear a distant explosion more 3 —— in summer than in 4 ——.

Answers
1 Air
2 Increases
3 Quickly
4 Winter

II. Alternate Response or True-False Items

The true-false test presents a series of statements or questions each of which is to be marked T (true) or F (false), right or wrong, yes or no, agree or disagree and the like. In each case there are only two possible answers.

Uses of Alternate Response Items : This type is used in measuring the ability to identify the correctness of statements of facts, definitions of terms, statements of principles and the like.

(i) *Ability to identify the correctness of statements* :

Read each of the following statements carefully and decide whether it is true or false. Draw a circle around the corresponding letter T or F.

T F A nephew is always younger than his uncle.
T F Nitrogen supports combustion.
Yes No Is 25% of 44 less than 12 ?
Yes No The weather forecast is always correct.

(ii) *Ability to distinguish fact from opinion* :

Read each of the following statements; if the statement is true, circle the T; if the statement is false, circle the F and if the statement is an opinion, circle the O.

T F O The earth is a planet.
T F O The earth revolves around the moon.
T F O There are no plants or animals on Mars.
T F O Mysore is the most beautiful city in India.

(iii) *Ability to recognize cause and effect relationship* :

This type of item usually contains two true propositions in one statement and the pupil is to judge whether the relationship between them is true or false.

Directions : In each of the following statements, both parts of the statement are true; you are to decide if the second part explains why the first part is true. If it does, circle the 'Yes' if it does not, circle the 'No'.

Yes No 1. Leaves are essential parts of a tree because they shade the trunk of the tree.
Yes No 2. Whales are mammals because they are large.
Yes No 3. Brenda and Rose are in the same class, because they are neighbours.
Yes No 4. Some plants do not need sunlight because they get their food from other plants.
Yes No 5. Water in the ocean evaporates because it contains salt.
Yes No 6. 408 is divisible by 4 because the last digit in the number is 8.

Merits : (1) It may be used with a wide variety of materials. (2) It can be scored easily and objectively. (3) It is the easiest objective type of item to construct. (4) It makes possible an extensive sampling of material in a relatively short space. (5) It is a time saver, thus allowing for frequent testing. (6) The directions are easily understood.

Limitations : (1) It is often ambiguous. (2) It is open to guessing and to chance effects. (3) Much subject matter cannot be stated as unequivocally *true* or *false*. (4) It may readily become a test of detatched and unrelated bits of information. (5) It may overstress rote memory at the expense of understanding.

(6) Items testing understandings, inferences, applications are usually hard to cast in this form.

A common criticism of this type is that a pupil may be able to recognize a false statement as incorrect but still not know what is correct. To overcome this difficulty some teachers ask the pupils to change all false statements to true. The limitations of this type of item are so serious that it would be wise to use this item only where other types are inappropriate for measuring the desired learning outcomes. This would include situations where there are only two possible alternatives such as *yes-no* or *right-wrong*.

Maxims for Alternate Response Items

1. *Double negative statements should be avoided* : Double negatives serve no useful purpose; they are hard to read and confusing.

Poor : T. F. One cannot interpret the results of tests which provide no norms.

Better : T. F. It is difficult to interpret the results of tests which do not provide any norms.

Under the tension of the test situation, an individual can easily skip over the *No* in the first example. This mistake in reading is unlikely in the latter example.

2. *Statements that are partly true and partly false should not be used* : Statements should be simple and direct.

T. F. Napoleon lost the battle of Waterloo because his army had insufficient ammunition.

T. F. Ice floats on water, because it is a solid.

In these examples one part of the statement is true and the other is false. In the second example the statement is true but the reason is false.

3. *Avoid tricky statements* : Tese are usually statements which appear to be true but which are really false, because of some inconspicuous word.

T. F. Water boils at 212°F and freezes at 32°C.

4. *Avoid specific determiners* : The meaning of a statement should be the only clue to the correct answer. Certain

determiners of a statement's truth or falsity are to be avoided. Strongly worded statements containing such words as *all, always, never, none, entirely, absolutely, totally* etc., have been found experimentally to be much more often false than true. On the contrary, sentences using the words *should, may, most, sometimes, often,* are more than apt to be true ones. Careful balance between true and false statements containing any of these will avoid their becoming specific determiners. Otherwise the test-wise student will use these cues to get credit for knowledge he does not possess. Still another clue seems to arise from the length of the statement. By actual count of words it has been shown that long sentences (with more than 20 words) are likely to be true.

Examples : Girls are *generally* more studious than boys.

All people are suggestible to a great extent.

Generally people working in the same profession have common interests.

All girls are more beautiful than boys.

The words in italics offer a clue to the pupils regarding the truth or falsity of the statements and the pupils get credit.

5. *Avoid ambiguous and indefinite terms of degree or amount* : Whenever possible, quantitative statements are better than such expressions as *frequently, more, in most cases, less* etc., as indicators of comparison. These latter terms are not interpreted in the same way by everyone who reads them.

6. *Approximately an equal number of true and false statements should be used* : It is not desirable to have a great imbalance of true and false statements.

7. *Random occurrence of true and false statements should be employed.*

8. *Authority must be mentioned for statements that might be considered controversial.*

9. *Use the simplest possible method of indicating the response* : Instead of asking the pupil to write *True* or *False*. Yes or No, let him write T or F or underline the correct response. One of the simplest ways is to write the symbol + for true and O for false in a blank space provided at the end of each statement.

10. *Correction for guessing should be used in scoring the test.*

III The Matching Item

This type is actually a special form of the multiple choice item. If differs from the multiple choice item in that instead of a single stem with a number of suggested responses, there are several problems whose answers must be drawn from a single list of possible answers. It is thus an economical form of including a number of multiple choice items in the same question. This test consists of two columns; each item in the first column is to be paired with a word or a phrase in the second column upon some basis suggested, e.g., dates may be matched with events, authors with books, substances with their properties, causes with effects, organs with their functions etc.

Merits : Many questions can be answered in a short space. It is a compact and efficient way of measuring the type of achievement such as matching terms and definitions, names and events etc. Its greatest usefulness comes in answering questions of the type *who*, *when*, *what*, and *where*. This type is capable of making a rapid survey of a particular phase of subject matter.

Limitations : It is not well adapted to the measurement of understanding as distinguished from rote memory. Items making up these tests should be homogeneous; it is difficult to get enough homogeneous items in the case of many of the outcomes such as interpretation of data, or application of principles, hence the use is restricted to test factual types of subject matter. Moreover the items are likely to include irrelevant clues to the correct answer.

Maxims for Matching Tests

1. *Indicate clearly the basis on which matching is to be done* : This should be specified both in the directions and in the column headings. Indicate also whether an answer choice may be used more than once.

2. *Check each list for unwarranted clues* : All items in the first group should agree in form and all items in the second column should likewise be in agreement e.g., if there are both

singular and plural forms of words in group A, it will be comparatively easy to spot out the corresponding response word from group B. Answers can thus be found by looking to grammatical details in the statements.

3. *Let the material be homogeneous* : If dates and events are mixed up with names of places and events or terms and definitions in the same group, there will be clues that encourage a process of elimination or a negative approach. Guessing is eliminated by the homogeneity of items.

4. *Include more responses in the right hand column than are needed* : Increasing the number of distractors makes the test difficult. Even some of the responses may be used more than once.

5. *The items in one of the columns should consist of single words or numbers or brief phrases* : They should be arranged in some logical order e.g., dates should be chronological, authors or terms should be alphabetical. This will shorten the working time.

6. *All items should be on the same page* : Turning the page back and forth in search of the desired responses is both confusing and time consuming.

7. *The number of items in a test should be from 5 to 10* : Too few pairs encourage guessing while too many increase the working time.

Following is an example of a poor exercise in matching items :

Directions : Place just one letter to the left of each number in column A.

Column A	Column B
——— 1. Terman and Merrill	A. Adult intelligence test
——— 2. Aptitude	B. Revised Stanford-Binet
——— 3. 1916	C. Innate capacity to learn
——— 4. Reliability and validity	D. Characteristics of a good test
——— 5. Wechsler	E. Stanford Revision of Binet

The above example violates several rules.

(i) There are no directions given to indicate the basis for matching. (ii) The items are not homogeneous; the pairs include author-test, term-definition, date-event; this lack of any common principle of association makes a negative approach fruitful. (iii) Many clues are present. Item no. 4 in both the columns is given in plural form. The two must to together. There is only one definition in column B, naturally it will go with the term *Aptitude* in column A. When four items are decided, the fifth response is automatically located, without having the knowledge about it. (iv) The arrangement of items does not follow any logical order.

In attempting this test the approach is negative. A clever student with a few hazy ideas about this subject can manage to fool the test. His score is perfect, but his thinking is cloudy, as he does not study this material systematically. In order to revise this exercise, all the above drawbacks must be removed.

Following are examples to show how the items should be prepared by observing the rules.

(i) *Example* : Each agricultural product on the left is an important export of one of the countries listed in colmn B. Place the letter of the country on the line before the product it exports.

	A		B
———	Coffee	a.	Argentine
———	Beef	b.	Australia
———	Wheat	c.	Brazil
———	Wool	d.	Canada
		e.	England

(ii) *Example* : In column B are words which illustrate a number of parts of speech; column A is a list of various parts of speech. Determine what part of speech a word from list B is and then identify it by putting its number before the proper item in column A.

	A		B
———	adjective	(1)	and
———	adverb	(2)	cat
———	noun	(3)	rapidly
———	preposition	(4)	jump
———	verb	(5)	from
		(6)	rich
		(7)	either
		(8)	their

The second example, though an easy one, will enable a teacher to locate grammatical confusions of children.

IV Master Matching Form

This is a variation of the matching type of item. This is an economical way of asking a number of multiple choice questions. It consists of a set of alternatives which serve as a master list and a number of statements or questions following it, each to serve as a stem. This pattern presents an efficient means of exploring the range of mastery of a concept.

Example : Suppose in a certain State Legislative Assembly, some members of the ruling party cross the floor and join the opposition party in passing a vote of censure against the State Ministry. Below are given a few statements that describe some of the after-effects of this event. Using your knowledge of the State administration, place in the bracket provided before each statement, corresponding alphabet as shown in the master list.

Master List

A—if it is most likely to happen
B—if it is remotely possible
C—if it is definite that it would not happen that way
D—if nothing can be said as to whether it would happen that way or not
E—if it is irrelevant

Statements

( ) 1. The speaker gives another opportunity to the cabinet to face a motion of confidence.
( ) 2. The Ministry hands over its resignation to the Governor.
( ) 3. The Governor accepts the resignation of the Ministry, but asks it to continue to be in office till a new Ministry is formed.
( ) 4. The speaker also resigns along with the Ministry.
( ) 5. The Governor of the State also changes with the Ministry.
( ) 6. The new Ministry presents its own budget to the Assembly for the remaining part of the year.
( ) 7. The Governor dissolves the Legislative Assembly.
( ) 8. The defeated party will bargain with opposition members to regain power.
( ) 9. The Chief Minister will go to a health resort for rest.

Uses : This form of question is very well suited for testing pupils' basic understanding in any field of knowledge by providing them a complex situation wherein they have to use or review a large number of related facts. It also tests pupils' ability to apply knowledge in known as well as in unfamiliar situations.

V Multiple Choice Items

A multiple-choice item consists of a problem and a list of suggested solutions. The problem may be stated in the form of a direct question or an incomplete statement and is called the *stem* or the *premise* of the item. The list of suggested solutions may include words, numbers, symbols or phrases and are called *alternatives.* The correct alternative is called the *answer* while the remaining alternatives are called *distractors,* whose function is to distract those pupils who are in doubt about the correct answer.

Uses of Multiple-Choice Items : This type is the most flexible and most effective of the objective type items. It is effective for measuring a variety of learning outcomes such as information, vocabulary, understandings, application of principles and ability to interpret data. Many standardized tests use this item exclusively because of its wide applicability. The following examples will show how the various learning outcomes can be measured by this type of item.

(a) *Measuring knowledge outcomes* : Knowledge of terminology, specific facts, principles and knowledge of methods.

(i) A small body that falls through the atmosphere towards the earth and is heated to incanscence as it falls, is called

(a) a comet (b) a meteor
(c) a planet (d) a satellite ... ( )

(ii) Practically all religions include a belief in

(a) one God (b) ten commandments
(c) immortality (d) saints. ... ( )

(iii) The mean of 5, 1, 4, 6 is :

(a) 4 (b) 5
(c) 6 (d) 3 ... ( )

(b) *Measuring outcomes at the understanding level* : Ability to apply facts and principles, ability to interpret cause and effect relationship and ability to justify methods and procedures.

(i) In the following sentence circle the word that makes the sentence correct.

whom
This is the boy which asked the question.
who

(ii) Chlorine is passed through drinking water before distribution in a city, because it is

(a) a bleaching agent (b) soluble in water
(c) an oxidising agent (d) a strong disinfectant. ... ( )

(iii) "If equals are added to equals, the results are equal," which one of the following is the best example of this statement?

(a) $1 + 8 = 3 + 6$ (b) $3 + 3 = 5 + 5$
(c) $2 + k = 2 + k$ (d) $4 + 0 = 4$ ... ( )

(iv) Which one of the following fractions is smaller than $\frac{1}{2}$?

(a) $\frac{3}{8}$ (b) $\frac{4}{6}$ (c) $\frac{2}{4}$ (d) $\frac{9}{16}$ ... ( )

(v) Which of these addition examples is represented by the shaded parts of the diagrams below? ... ( )

(a) $\frac{1}{2} + \frac{1}{3}$

(b) $\frac{1}{4} + \frac{1}{3}$

(c) $\frac{2}{3} + \frac{3}{4}$

(d) $\frac{1}{4} + \frac{2}{3}$

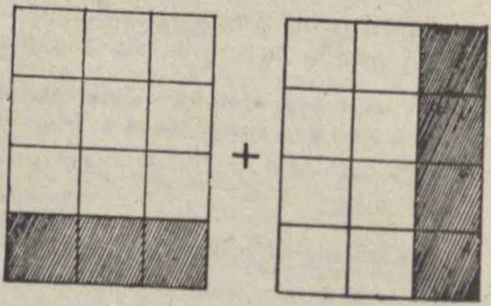


Fig. 11.1 : A problem on addition.

(c) *Measuring outcomes at the application level* : This type can be used for testing students' ability to justify methods and procedures.

Directions : This is a test to see how well children can do number problems.

Problem : Last month Mukesh and his older sister Nanda helped their father in the factory on every Saturday.

Mukesh worked for 16 hours in the month at Rs. 3.50 an hour while Nanda worked for 12 hours in the month at Rs. 2.25 an hour.

1. We are told in the story that
   a. Nanda worked for 16 hours
   b. Mukesh was paid at Rs. 2.25 an hour
   c. Mukesh worked for 16 hours
   d. None of these ... ( )
2. What could Nanda do to find out how much money she has earned ?
   a. Add 12 hours and 16 hours
   b. Add Rs. 3.50 and Rs. 2.25
   c. Multiply 16 by 3.50
   d. Multiply 12 by 2.25 ... ( )
3. For finding out the total amount earned by Mukesh and Nanda the following procedure is essential
   a. (16 × 12) × 3.50
   b. (16 × 3.50) × (12 × 2.25)
   c. (3.50 + 2.25) × (16 + 12)
   d. (16 × 2.25) × (12 × 3.50) ... ( )
4. Nanda earned
   a. Rs. 24
   b. Rs. 42
   c. Rs. 36
   d. None of these ...( )
5. Mukesh earned
   a. Rs. 56
   b. Rs. 27
   c. Rs. 36
   d. Rs. 42 ... ( )
6. How much more did Mukesh earn than Nanda?
   a. Rs. 29
   b. No more
   c. Rs. 6
   d. None of these ...( )

Such exercises are also given in reading comprehension tests (Page 129). Various examples of objective based multiple choice items are given in Appendices I and II of this book.

Merits : (i) This type can measure a variety of learning outcomes such as knowledge, understandings and applications. (ii) It is free from ambiguity and vagueness which are present in the completion type item. (iii) Guessing is minimized as the number of alternatives is increased to four or five. The item thus becomes more reliable. (iv) Distractors selected by the pupils provide clues to the teacher for diagnosis of pupils' difficulties. (v) This type need not meet the absolute standard of truth and falsehood demanded of true-false type of item. The pupil is to select the best alternative from among the given ones. This makes it possible to measure learning outcomes in the numerous subject matter areas.

Limitations : (i) This type is not well adapted for measuring the ability to organize and present ideas. (ii) It is difficult to locate sufficient number of incorrect but plausible distractors. Plausible alternatives are hard to find.

Maxims for Multiple Choice Items

(1) *Beware of irrelevant grammatical cues* :

Poor : A man who designs houses is called an
(i) contractor, (ii) architect, (iii). carpenter, (iv) engineer.

Better : A man who designs houses is called
(i) a contractor (ii) an architect, (iii) a carpenter, (iv) an engineer.

In the first example the article ' an ' in the stem automatically eliminates two of the response options. The item then becomes a true-false item. Similarly the distractors must be grammatically consistent with the stem such as the plural or singular form or tense of the verb etc.

(2) *Make all responses plausible* : The difficulty of an item will depend upon the closeness of the distractors. The distractors must be framed so as to tempt the student who has only superficial knowledge of the point involved. If the distractors are not plausible, a student will get the correct answer just by eliminating the absurd ones.

Poor : An example of sublimation is
(i) binocular disparity, (ii) inhibition,
(iii) creative art, (iv) neurosis. ... ( )

Better : An example of sublimation is
(i) denying expression, (ii) creative art,
(iii) direct fulfilment, (iv) expressing difference of opinion ... ( )

In the first example the distractors are so implausible that they can be eliminated by any one. To ensure the validity of an item the distractors must be closely related to the problem. In the improved version there are no clues and a negative approach of eliminating wrong choices will not work.

(3) *The correct answer should not be consistently longer or shorter than the distractors* : There is a tendency for the correct answer to be longer than incorrect answers, because of qualifying phrases that make it true.

(4) *Answers should follow a random pattern* : Students are likely to detect any regularly recurring pattern in the sequence of responses.

(5) *The stem of the item should clearly formulate a problem* :

Poor : The temperature of a place
(i) is high if it is nearer the equator,
(ii) is low if it is at lower altitudes,
(iii) is high if it is nearer the sea,
(iv) is high if it is on higher latitude ... ( )

Better : The temperature of a place mainly depends upon
(i) the soil,
(ii) the supply of water,
(iii) the latitude,
(iv) the altitude from sea-level. ... ( )

In the first example the premise does not set forth any definite problem and the item then becomes a series of unrelated true-false items dealing with some general topic.

(6) *Wording should be as concise as possible* : Irrelevant material should not hide the crucial aspect of the problem.

Poor : Lata, a ten year old girl who is studying in standard V in a progressive city school, has an IQ score of 105 on the Stanford-Binet. This would classify her as :
(i) Superior, (ii) Dull, (iii) Average, (iv) Border-line.

This item can be improved by eliminating the useless facts that have unduly loaded the stem. When however the purpose of the test is to judge the ability of the pupil to identify and pick out essential facts from a mass of data, the item may be loaded even with irrelevant material.

(7) *Be sure that there is one and only one correct or clearly best answer.*

(8) *Each item should have four or five choices* : Increasing the number of choices reduces the amount of guessing. Use the 'correction for chance' formula while scoring the test.

VI The Interpretive Type Item

This type of item consists of an introductory selection of material giving the necessary background and setting the problem, followed by a series of questions asking for interpretation of the material. This material may be in the form of passages, graphs, tables, maps, charts etc.

We can measure various learning outcomes with the help of interpretive type. These learning outcomes are based on the higher mental processes for example :

(i) ability to interpret relationship
(ii) ability to recognize relevance of information
(iii) ability to formulate and recognize valid conclusions
(iv) ability to recognize assumptions underlying conclusions
(v) ability to recognize the limitations of data
(vi) ability to apply a princple.

We shall consider how the interpretive type can be used to measure some of these outcomes.

(a) *Ability to recognize the relevance of information* : The following example is meant for elementary school pupils.

" Bill lost his overshoe on the way to school. He wanted to put a notice on the bulletin board so that other children could help him find it. Which one of the following sentences tell something that would help children find the overshoe ? "

Circle *Yes* if it would help, circle *No* if it would not help.

Yes No 1. The overshoe was red
Yes No 2. It was very warm
Yes No 3. It was for his right foot
Yes No 4. It was his birth-day present
Yes No 5. It was nice looking
Yes No 6. It had a zipper

(b) *Ability to interpret data* : The attendance of children in a school for one week is given below. The total number of pupils is 400. Read the following graph and answer the questions given below.

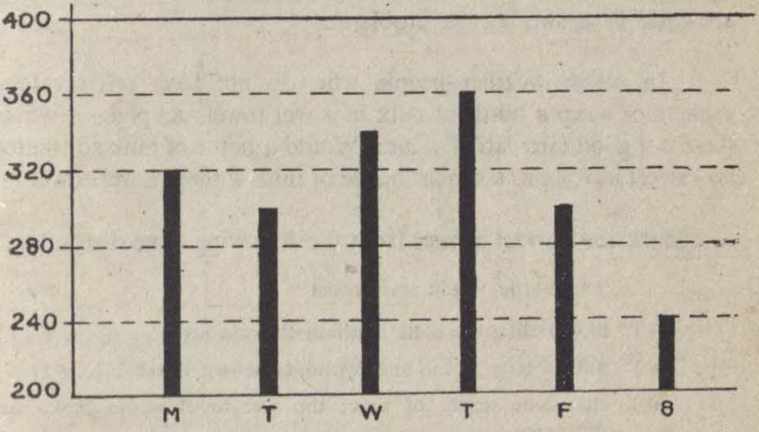


Fig. 11.2 : Interpretive type of item.

(1) On which day was the highest attendance ? ... ( )
(2) What was the attendance on Tuesday ? ... ( )
(3) On which day was the lowest attendance ? ... ( )
(4) What was the percentage of attendance on Friday ? ... ( )
(5) What was the average attendance for the whole week ? ... ( )

(c) *Recognizing the limitations of data* : A given statement is followed by several proposed conclusions. You are to judge whether each of the proposed conclusions logically follows beyond a reasonable doubt from the given information.

"A report of the U.S. census states that during the year 1940, there were approximately 1,656,000 marriages and 264,000 divorce granted in the United States."

Conclusions :

(i) Getting a divorce is a quick and easy matter in the U.S.
(ii) If the above ratio still holds true, then about six times as many people get married each year as get divorced.
(iii) The divorce rate in the United States is much too high.

(*Answer* : Only the second conclusion necessarily follows; the other two do not.)

(d) *Ability to formulate and recognize valid conclusions* : You are given a problem. Below it are given a few statements which can be used to answer the problem. Check the ones that are valid in answering the problem.

" In warm weather people who do not have refrigerators. sometimes wrap a bottle of milk in a wet towel and place it where there is a good circulation of air. Would a bottle of milk so treated stay sweet as long as a similar bottle of milk without a wet towel ? "

Mark the correct answer from the following three distractors :

The bottle would stay sweet

(i) longer than the bottle without the wet towel ... ( )

(ii) not as long as the one, without the wet towel ... ( )

(iii) the same length of time; the wet towel would make no difference ... ( )

Check the statements below which give the reason or reasons for your explanation above.

(1) Thunderstorms hasten the souring of milk.

(2) The souring of milk is the result of the growth of bacteria.

(3) Wrapping the bottle prevents bacteria from getting into the milk.

(4) A wet towel could not interfere with the growth of bacteria in the milk.

(5) Wrapping keeps out the air and hinders bacterial growth.

(6) Evaporation is accompanied by an absorption of heat.

(7) Bacteria do not grow so rapidly when temperatures are kept low.

(8) Milkmen often advise housewives to wrap bottles in wet towels.

(Statements 2, 6 and 7 give the right principle. No. 1 is mere a superstition. No. 8 is based on authority, the rest are wrong.)

(e) *Ability to recognize warranted and unwarranted generalizations* : In the interpretation of data, pupils should be able to determine which conclusions are supported by the data, which are refuted by the data and which are neither supported nor refuted by the data.

Example

Mortality of persons from motor vehicle accidents in the United States 1957-58.

Age-Group	Death rate per 100,000	
	Males	Females
1-4	10.5	8.0
5-14	10.4	5.4
15-19	54.2	16.4
20-24	76.3	12.7
25-44	35.6	9.1
45-64	33.1	12.9
65 and over	58.4	22.5
All ages	32.9	11.1

The following statements refer to the data in the above table. Read each statement and mark your answer according to the following key.

Circle : S if the statement is supported by the data in the table.
R if the statement is refused by the data.
N if the statement is neither supported nor refuted by the data.

S R N 1. The death rate from motor vehicle accidents is higher for men than for women.
S R N 2. Motor vehicle accidents are a major cause of death among young men between the ages 20-24.
S R N 3. Men over 65 years of age drive no more safely than teen age boys between 15 and 19 years of age.
S R N 4. The largest number of people killed in motor vehicle accidents are 65 years of age or over.
S R N 5. When all ages are combined, only about 11 percent of female deaths can be attributed to motor vehicle accidents.
S R N 6. Girls are not so reckless as boys in driving a car.

(*Key* : S, N, N, R, R, N.)

(f) *Ability to identify valid cause and effect relationship* : Below are listed pairs of events. In the blank space between the two events, write :

A. If the first event is the cause of the second event.
B. If the first event is the result of the second event.
C. If the two events have no cause and effect relationship.
D. If both events are results of the same general cause.

First Event		Second Event
1. Clouds are moving in the sky	——	Heavy dust is in the air
2. A weight is placed on the end of a suspended spring	——	The spring stretches
3. A rising column of air was cooled	——	A cloud is formed
4. Price of gold shoots up	——	Rupee is devaluated
5. The branches of a tree move to and fro	——	A nearby windwill turns
6. Air contains oxygen	——	Air contains nitrogen

(g) *Deductions* : Two statements are given on the left. Tick the one statement on the right, which must necessarily follow from them.

(a) Meena is not as old as Sujata

(b) Sujata is older than Vimal

Therefore it follows that :

(i) Meena is the youngest

(ii) Vimal is the youngest

(iii) Meena is the eldest

(iv) Sujata is the oldest

(h) *Ability to interpret relationship* : Pictorial materials in the form of graphs, maps, diagrams serve a useful purpose in interpretive exercises. We have already seen one example in Fig. 11.2. Here is another type.

Mangesh, a high school teacher spends his monthly earnings on various items as shown in Fig. 11.3.

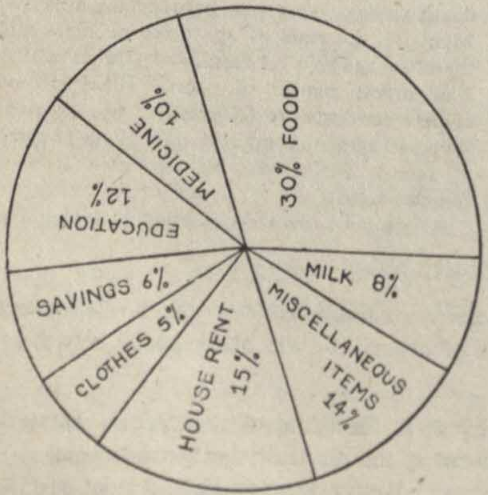


Fig 11.3 : Interpretive type item.

Answer the following questions :

1. The graph shows that Mangesh spends the largest amount on
   - a. House rent
   - b. Education
   - c. Food
   - d. Miscellaneous ... ( )
2. Expenditure on house rent and clothing together is the same as the expenditure on
   - a. Milk and education
   - b. Medicine and Milk
   - c. Miscellaneous items
   - d. None of these. ... ( )
3. His savings would be doubled if his expenditure is reduced by
   - a. 8%
   - b. 4%
   - c. 12%
   - d. 6% ... ( )
4. The expenditure on food is two times the expenditure on
   - a. Miscellaneous items
   - b. House rent
   - c. Education and Milk
   - d. Medicine and Milk. ... ( )

(i) *Ability to recognize the sequence of events* : A number of events are given in random order. Students are to arrange them in proper order.

1. The picture 'Guide' had already started ... ( )
2. Everything went wrong last evening ... ( )
3. *Miss Vaheeda* wanted to leave the office early ... ( )
4. By the time she reached the bus stop, there was a long queue ... ( )
5. The picture was excellent ... ( )
6. But there was so much work that she left after six ... ( )
7. She reached the cinema a little after seven o'clock ... ( )
8. When she caught the bus, she had to stand all the way. ... ( )

(j) *Ability to solve a given problem* :

If L = 2, M = 3, N = 4, P = 10

(a) Which two letters add up to 12 ? ... ( )

(b) Which two letters multiplied give 30 ? ... ( )

(c) Which letter gives the answer when M is multiplied by N and then P is substracted from the product ? ... ( )

(d) Which letters multiplied together give the same answer as P and L added together ? ... ( )

Merits : (i) The interpretive exercise makes it possible to measure more complex learning outcomes than can be measured with the single objective item. Pupils can demonstrate thinking and problem solving skills. (ii) The use of reference material and development of self-study techniques are encouraged because of the interpretive exercise. (iii) Because of the introductory material provided, the influence of irrelevant factual information on the measurement of complex learning outcomes is minimised.

Limitations : (i) Selecting interpretive material which is new to the pupils but which is related to the objectives of instruction, is somewhat difficult. (ii) This type places heavy demand on reading skill. The poor reader is handicapped by the difficulty of the reading material. (iii) This type cannot measure a pupil's over-all approach to problem solving such as defining a problem, formulating hypothesis, organizing data and drawing conclusions. In this respect it does not favourably compare the essay test. It measures only specific aspects of the problem solving skill.

FOR FURTHER STUDY

1. Illustrate the use of multiple choice items in the measurement of various learning outcomes. What are the merits and limitations of this type ? State six maxims for constructing the multiple choice item.
2. In what respects is the interpretive type better than other objective items ? Give two examples to show the use of this type of item.
3. What are the merits and limitations of :
   (i) True-False items and (ii) Matching items ?
   State with examples the maxims for constructing any one of these types.
4. The following T-F items are constructed by violating some of the maxims. State in each case the maxim that is violated.
   T.F. (1) It is sometimes permissible to administer a test orally rather than in writing.
   T.F. (2) Well constructed test items cannot have but one interpretation.
   T.F. (3) "Don't put all your eggs in one basket" is a good rule in test construction.
   T.F. (4) The traditional examination should be relegated with celerity to the limbo of anachronistic entity.
   T.F. (5) Statements containing the words all, always, often, sometimes, are more likely to be false than true.

[For additional objective based test-items refer to items given in the appendix.]

Answers :

Correct order of events in exercise (i) on page 227 :
2, 3, 6, 4, 8, 7, 1, 5.

* * *

12

TOOLS OF EVALUATION : III

[Qualitative Techniques]

We have seen in the two previous chapters that a student's performance can be evaluated with the help of class-room tests which give a quantitative measure of pupil behaviour. However we do not get answers to such questions as :

(i) Is Jyoti sociable ?
(ii) Does she show initiative in group activities ?
(iii) Can she pursuade others to her view points ?
(iv) Does she take interest in library reading ?
(v) Does she have a consideration for others ?
(vi) Are her feelings easily hurt ?
(vii) What does she feel about her teachers ?

Evaluation based merely on quantitative measures cannot be considered valid, since it neglects the qualitative aspect of pupil behaviour. Qualitative measures can further be classified as :

(i) Observational Techniques (ii) Self-report Techniques
(iii) Projective Techniques (iv) Sociometric Techniques.

I. Observational Techniques

A teacher can observe her pupils in a number of situations in the class-room, on the play ground, in a social gathering, in the laboratory etc. She can use a number of devices such as a rating scale, a check-list and an anecdotal record. We shall now critically examine these devices.

1. Rating Scales

Rating scale is a device for obtaining judgments of the degree to which an individual possesses certain behaviour traits and attributes not readily detectable by objective tests. The rating of one person by others is among the oldest of practices. For the evaluation of an individual, rating scales are used by teachers, employers, colleagues and others who have had sufficient contact with the person in question. Rating scales may be devised for evaluating a great variety of traits such as tact, generosity, leadership, co-operativeness, resourcefulness, emotional control and many others. Typically a rating scale consists of a set of characteristics or qualities to be judged and some type of scale for indicating the degree to which each attribute is present. Rating scales are also used for evaluating *procedures* such as playing on an instrument, singing, working in the laboratory etc., and for evaluating *products* such as typing, cooking, woodwork, handwriting etc. Rating scales are generally used by officers for writing the confidential report of their subordinates. Headmasters use the rating scale for estimating the qualities of teachers such as initiative, co-operativeness, resourcefulness, class control etc.

What are the types of Rating Scales ?

(*a*) *Numerical Rating Scale* : One of the simplest types of rating scales is that where the rater checks or circles a number to indicate the degree to which a characteristic is present. Typically, each of a series of numbers is given a verbal description which remains constant from one characteristic to another.

Example

Directions : Indicate the degree to which this pupil contributes to class discussions by encircling the appropriate number. The numbers represent the following values :

5—Outstanding, 4—above average, 3—average, 2—below average and 1—unsatifactory.

(i) To what extent does the pupil participate in discussions ?
.. .. .. 1 2 3 4 5

(ii) To what extent are his comments related to the topic under discussion ?
.. .. .. 1 2 3 4 5

The numerical rating scale is useful when the characteristics or qualities to be rated can be classified into a limited number of categories and when there is general agreement concerning the categories represented by ach number.

(*b*) *Graphic Rating Scale* : The distinguishing feature of the graphic rating scale is that each characteristic is followed by a horizontal line. The rater's evaluation is indicated by placing a check or cross on a line to indicate presence or absence of a given trait. Various categories are marked on the line, but the rater is free to check between these points if he so desires.

Example

Directions : Indicate the degree to which this pupil shows initiative and leadership in doing his work by placing an × anywhere along the horizontal line under each item.

(i) To what extent does he show initiative ?

never	seldom	occasionally	frequently	always

(ii) To what extent does he show leadership ?

never	seldom	occasionally	frequently	always

The use of single words such as *never, seldom* etc., to identify the various categories has no great advantage over the use of numbers used in the previous type. There is little agreement among raters concerning the meaning of such terms as *seldom, occasionally* and *frequently*. What is needed is behaviour description which indicates more specifically the presence or absence of a trait.

(*c*) *Descriptive Graphic Rating Scale* : This rating form uses descriptive phrases to identify the points on a graphic scale. The descriptions convey in behavioural terms what the individuals are like at different steps along the scale.

Example

Direction : Mark your ratings on each of the following characteristics by placing an X anywhere along the horizontal line, under each item. In the space for comments, include anything that helps clarify your rating.

(i) To what extent does the pupil participate in discussions ?

never participates, quiet, passive	participates when specifically asked to do	participates as much as other members	participates on many occasions	participates more than any group member

Comments : ....

(ii) To what extent does he show leadership?

probably unable to lead others	lets others take a lead	sometimes leads in minor affairs	sometimes leads in important affairs	marked ability to lead others

Comments : ....

This scale is generally the most satisfactory for school use. It clarifies to both the teacher and the pupil the types of behaviour that represent different degrees of progress. Such specific descriptions contribute to greater objectivity of the rating process.

(*d*) *Ranking Methods* : Some rating procedures do not require a printed scale. Probably the most applicable and best known of these is the simple rank-order method. With this approach, the pupils (or products) being rated are merely ranked in the order in which the rater estimates they possess the characteristics being judged. Ratings made by a number of judges may then be combined into a composite or average rank. A practical modification of the usual ranking method requires the rater to sort the pupils (or products) into a given number of groups on the basis of over-all quality and then letter grades such as A, B, C, D and E are assigned to these groups.

Common Errors in Rating

Certain types of errors occur so frequently and persistently in ratings that special efforts are needed to counteract their influence. These include errors due to (i) personal bias, (ii) halo effect and (iii) logical error.

(i) Personal bias errors are indicated by a general tendency to rate all individuals at approximately the same position on the scale. Some rates tend to use the high end of the scale only. This is generally the most common type of bias and is referred to as the *generosity error*. There are a few who prefer the lower end of the scale. This is known as the *severity error*. A third type is shown by the rater who avoids both ends of the scale and tends to rate everyone average. This is called the *central tendency error*. Of all the three the generosity error is the one that occurs most frequently.

(ii) The *halo* effect is an error that occurs when a rater's general impression about a student influences his ratings on various traits. If the rater has a favourable attitude toward the person being rated, he will tend to rate him high on all traits; if his attitude is unfavourable he will tend to rate him low. This differs from the generosity and severity errors where the rater tends to rate everyone high or everyone low.

(iii) A *logical* error results when the rater has some beliefs concerning the relationship of two characteristics. For example, teachers who hold the common but false belief that gifted pupils have poor social adjustment will tend to underrate them on social characteristics. The errors here, do not result from biases toward certain pupils but rather from the rater's preconceived notions concerning the human nature.

What are the requirements of a good rating scale ? : (1) Traits must be clearly defined. This is essential so that the trait may be clearly and uniformly understood by all judges. Many of the errors in rating arise from the use of general and vague terms such as initiative, leadership etc. Brief discriptions used with the descriptive graphic rating scales would help overcome this weakness. (2) Number of divisions on the scale should be neither too numerous nor too few. Optimum number of divisions is perhaps five to seven. Raters should be permitted to mark at intermediate points. (3) The rater may be asked to quote instances in support of his judgment. This will put a check to careless rating on the part of the rater. (4) Raters should be instructed to omit ratings of characteristics, if they had little or no opportunity to observe the traits. (5) The rater should be instructed to avoid the generosity error, the logical error and the error due to halo effect. (6) Ratings from several observers should be combined wherever possible. Pooled ratings of several teachers will generally provide a more reliable description of pupil behaviour than that obtained from any one teacher.

Validity and Reliability of a Rating Scale : Ratings for intelligence and for special aptitudes can be validated against objective test scores, but ratings of personality traits cannot be so validated, since no objective test scores are available for personality traits. Rating is many times affected by our prejudices for or against an individual. A student who is submissive, who salutes the teacher everyday is likely to be rated more favourably by the teacher than a student who is bright but may not be so submissive. Validity is lowered in such cases. Hence we have to depend on the consensus of judges. If three or more judges agree that Dinesh is a skilled worker or a friendly person, the average of these ratings is reliable (consistent) and valid (worthy of confidence). If two supervisors decide independently that Miss Kaveri has emotional stability and shows a sense of responsibility in her school work, our confidence in these two judgments is greater than if only one supervisor has so stated. In general reliability and validity increase with the number of agreements when ratings are made independently.

2. Anecdotal Records

Teacher's daily observations provide them with a wealth of information concerning the learning and development of their pupils. For example, a chemistry teacher notices during a laboratory period that Ramesh is slow and inefficient in setting up his equipment, that Madhav finishes his experiments early and that Hemant handles the chemicals in a careless and dangerous manner despite repeated warnings. Such daily incidents and events have special evaluative significance. They enable us to determine how a pupil typically performs or behaves in a variety of situations. A simple and convenient method of recording such observations is provided by *anecdotal records.* An anecdotal record is a factual description of *meaningful events* which a teacher has observed in the lives of her pupils. It is a verval *snap-shot* which represents the pupil's behaviour. Each event is described shortly after it happens. A good anecdotal record consists of an objective description of an incident and includes an interpretation of the observed behaviour.

1. Example : "I have found Sheilla day after day in the school library reading magazines, yet she never has time to correct her papers or to work carefully on her English assignments."

Comments : She likes to read but not to work on her school work.

2. Example : Judy is seen sitting in a corner when other girls are playing during the recess.

Comments : She finds it difficult to mix with other girls in her class. She should not be allowed to remain aloof like that.

Advantages of Anecdotal Records : (i) The most important advantage is that these records provide a description of actual behaviour in natural situations. These records provide a check on other evaluation methods. (ii) These records can be used with very young pupils and with others who are retarded in basic communication skills. They are especially valuable here because other techniques are of limited use in such cases.

Limitations : (i) It is a time consuming task. (ii) It is difficult to get an objective record of pupil behaviour because the teacher's own biases and preconceived notions enter into his observations, for example, a teacher will tend to notice more desirable qualities in those pupils he likes best, and more undesirable qualities in those he likes least. Training in observation and reporting can reduce such distortions to a minimum. (iii) Everyone's behaviour fluctuates some what from situation to situation, for example, Dileep appears interested in a mathematics class bus disinterested in an English class. Therefore we need to observe him over a period of time and in a variety of situations.

Precautions in writing Anecdotal Records : (1) Do not pick out an incident out of context as it is difficult to interpret behaviour, apart from the situation in which it occurred. (2) Make a record soon after the incident is observed, otherwise, the observations become warped and distorted with the lapse of time. (3) Limit each anecdote to a brief description of a single specific incident. (4) Keep the factual description of the incident and your interpretation of it separate. The description of an incident should be as accurate and objective as you can make it. While writing, avoid such terms as *lazy*, *unhappy*, *shy*, *hostile*, *ambitious* etc., since such terms reflect your views. You may use such words when you interpret the behaviour. (5) Record both positive and negative behavioural incidents. There is a general tendency for teachers to note more readily those behaviours which disturb them personally and which interfere with the smooth processes in the class room. For evaluation purposes it is equally important to record the positive aspect of a student's behaviour. (6) Collect a number of anecdotes on a pupil before drawing inferences concerning the typical behaviour. A single behavioural incident is seldom very meaningful in understanding a pupil's behaviour.

3. Check Lists

A check list is similar in appearance and use to the rating scale. The basic difference between them is in the type of judgment called for. A rating scale provides an opportunity to indicate the degree to which a characteristic is present. The check list, on the other hand, calls for a simple 'yes-no' judgment. It is basically a method of recording whether a characteristic is present or absent. How does Bob handle the tools in a workshop? How does Barbara play on the piano? How does Meena conduct the class? We can evaluate such varied skills with the help of a check list. The check list helps us to know whether an individual possesses the characteristic. It does not help to know the degree to which the characteristic is present. Obviously a check list should not be used where degree or frequency of occurrence are important aspects of the appraisal.

In making a check list, the teacher must first decide what kinds of behaviour are important enough to be recorded. A list of these characteristics is combined with a list of pupils' names and the check list is ready for use.

Check lists are used by some schools in their report cards. (Fig. 9.2, page 175). Teachers try to break down the broad skill or content areas and vague personality traits into a large number of more limited and specific components. Consider a check list used to report the progress in languages (Fig. 12.1). In using the check list, all that the teacher needs to do, is to make a tick mark in the columns. By breaking down the broader skill into its components, we try to make the measuring device objective.

While observing the lessons of student-teachers, the method-masters should use a check list to denote the characteristics of teaching. A lesson can be objectively evaluated with the help of a check list. Such a list can be prepared by selecting some components as described on pages 33 and 79.

Language Skills	Standard....

Check each child 2 to 3 times during a month to determins if progress has taken place.

Behaviour to be observed	Names of Children				
	1	2	3	4	etc.
1. Understands what he reads					
2. Reads well aloud					
3. Takes part in discussions					
4. Expresses himself in writing					
5. Writes legibly					
6. Uses correct spellings in writing					

Fig. 12.1—A checklist for evaluating language skills

4. Cumulative Records

Evaluation should be based on an accumulation of different kinds of evidence. The record for each student should contain all the significant information gathered through various evaluation tools. The cumulative record is itself an important tool.

Teachers in secondary schools usually see only a segment of each pupil. They know him chiefly as a student of English or mathematics or science but not as an individual. The marks given by teachers in their respective subjects constitute only isolated bits of information and by themselves do not provide an adequate basis for understanding the individual student.

A cumulative record is a permanent record of a student which is kept up to date by the school. It is his educational history with information about his school achievement, attendance, health, test scores and similar pertinent data. Evaluation of a student's status at any particular time is made easier and richer by such a cumulative record.

II. Self-Report Techniques

We shall now consider evaluative tools which are classified as *Self-Report Techniques.* Every individual has a wealth of information about himself. Some of this information such as how he feels about certin situations, what activities interest him most, can be obtained directly from the individual. For evaluation purposes, this information is usually obtained by *interviews* and *questionnaries.*

1. Interview

Of the various methods of obtaining information directly from an individual, the oldest and the best known is that of the personal interview.

Merits: The face to face contact provided by the interview gives it several advantages as a self-resport procedure. First, considerable flexibility is provided. The interviewer can clarify his questions if they are not readily understood. He can skim over certain areas and probe intensively in others. Second the interviewer can carefully observe the interviewee during the session, noting the topics on which the subject is evasive and the areas in which he is most expansive.

Limitations : (i) This technique is extremely time consuming. (ii) The flexibility that is seen in this tool, destroys the comparability of information from one interviewer to another and from one interviewee to the next. It makes it possible for an interviewer to ride a personal hobby and ignore many obvious areas of inquiry. (iii) Since the utility of this tool depends on the willingness of the individual to give honest answers, it is decidedly limited for purposes of evaluating pupil progress.

Uses : A teacher can get much useful information regarding the difficulties of pupils in the school-life. Brenda finds it difficult to adjust with other girls. Kumud lags behind in the mathematics class. The teacher can diagnose the difficulties of these students through interview. This technique is also used in guidance and counselling.

2. Inventories or Questionnaires

The questionnaire method of obtaining information from individuals is most commonly used in systematic attempts to evaluate interests, attitudes and other aspects of personal and social adjustments. It has an advantage over the interview method in that it can be used in group situations. Each individual is presented with a series of questions or statements to which he must respond by answering *yes* or *no*, *agree* or *disagree* or in some similar manner indicate his feelings and opinions. Such self-report techniques are generally referred to as *inventories.*

An inventory consists of a standard set of questions pertaining to some particular area of behaviour, administrated and scored under standard conditions. It is a sort of a standardized, written interview which makes possible the collection of a large amount of information quickly. These inventories are sometimes referred to as *personality tests*, but this is a misnomer. The inventories do not call upon a pupil to perform at his best or maximum level; rather they depend upon his willingness and ability to describe what is typical, or generally true of himself. Hence the word personality *inventory* should be used instead of the word *test.*

(a) Personality Inventories

The personality inventory deals with motives and needs as well as with emotional and social factors. Adjustment or maladjustment to life is revealed by a person's self-report of his worries, fears, feelings of insecurity or of depression, frustration, lack of confidence and the like. There are many standardized personality inventories; we may consider a few items from the well known *Bell's Adjustment Inventory.*

The inventory consists of questions to be answered *Yes, No* or ?. It has been designed to estimate personal adjustment in four areas: home adjustment, health adjustment, social adjustment and emotional adjustment. There are no *right* or *wrong* answers. The individual is to indicate his answer by drawing a circle around *Yes, No, ?* There is no time limit.

Yes No ? (1) Do you daydream frequently ?
Yes No ? (2) Do you enjoy social gatherings just to be with people ?
Yes No ? (3) Do you ever have a strong desire to run away from home ?
Yes No ? (4) Do you find it easy to ask others for help ?
Yes No ? (5) Are your feelings easily hurt ?
Yes No ? (6) Do you feel tired most of the time ?
Yes No ? (7) Do you feel that your parents have been unduly strict with you ?
Yes No ? (8) Are you often in a state of excitement ?
Yes No ? (9) Have you felt that your friends have had a happier home life than you ?
Yes No ? (10) Do you have difficulty in getting rid of a cold ?
Yes No ? (11) Does criticism disturb you greatly ?
Yes No ? (12) Do you find it easy to make friendly contacts with members of the opposite sex ?
Yes No ? (13) Are you often sorry for the things you do ?
Yes No ? (14) Are you subject to attacks of indigestion ?
Yes No ? (15) Are you bothered by the feeling that people are reading your thoughts ?
Yes No ? (16) In school is it difficult for you to give an oral report before the class ?

Limitations of Personality Inventories : (1) The replies to various questions can be easily faked. An individual may try to give a picture of good adjustment by false answers. (2) Even if pupils want to be frank and honest, they may lack sufficient insight into themselves to be able to give objectively true responses. For example, pupils may simply not know the truth about themselves concerning such questions as : " Are you often sorry for the things you do " (3) The terms *often, frequently, sometimes* etc., do not mean the same things to different individuals. These limitations lower the validity of the inventories.

Uses : (1) These inventories are probably most useful as a screening instrument for identifying pupils who should be studied more closely by the school counsellor. The inventory may help locate the pupils who are poorly adjusted and unhappy and who need guidance in making proper personal adjustment. (2) Such inventories stimulate the pupils to evaluate critically their own personality characteristics.

Dr. (Miss) M. Bengalee of *St. Xavier's Institute of Education, Bombay* has standardized an inventory known as *Youth Adjustment Analyser*. This is a personality inventory measuring an individual's personal and social adjustment through five scales namely Parental attitude, Home adjustment, Aggression, Neuroticism and Clinical interest. It includes a lie scale which can detect dishonest responses in the case of individuals who try to avoid correct responses. It is available in four languages namely, English, Hindi, Gujerati and Marathi. Standardization was carried out on a sample of students from Maharashtra and Gujerat.

The scale has been validated against external criteria. The scale possesses both high reliability and validity. Norms are given in the form of categories ranging from excellent to very unsatisfactory adjustment for both school and college students.

The author has taken great pains in bringing out this excellent in the form of adjustment analyzer. The scale has proved its merits and it is one of the widely used inventories.

(b) **Interest Inventories**

There are several methods of obtaining information concerning a pupil's interests. Through direct observation we can note which areas of study receive his greatest attention, what type of books he reads and which activities he selects in his leisure time. Through direct questioning we can have the pupil tell or write about the things he would most like to do. These are informal methods and they have their obvious limitations. Answers to direct questions about interests are often unreliable and superficial

Standardized interest inventories overcome the limitations of informal methods. They have been designed primarily for use in educational and vocational guidance. We may consider one of the widely used interest inventories.

Kuder Preference Record : The major purpose of this inventory is to indicate relative interests in a small number of broad areas rather than in specific occupations. The items are arranged in groups of three. In each set of three, the examinee is asked to indicate which activity he would like most and which he would like least.

P. Visit an art gallery	○ P ○
Q. Browse in a library	○ Q ○
R. Visit a museum	○ R ○
S. Collect autographs	○ S ○
T. Collect coins	○ T ○
U. Collect butterflies	○ U ○

There are two circles on the line against each activity. The circle on the left is to be marked when the activity is liked most, and the circle on the right is to be marked when the activity is liked the least. The items cover a wide range of activities which are scored to yield an interest profile in ten areas namely outdoor, mechanical, computational, scientific, persuasive, artistic, literary, musical, social service and clerical. This inventory is an excellent measure of the range of expressed interests and as such is valuable in educational and vocational guidance.

(c) Attitude Scales

When we know that a man is a socialist, or a communist or an atheist, we feel fairly sure that we can predict his answers to questions dealing with politics or religion. An attitude is a consistent point of view, a way of behaving toward an institution, a social group, or toward personal, political or religious issues or practices. Attitudes may be fairly narrow or quite broad, and they may be strongly or weakly held. In the school situation, a teacher has to look to the development of desirable social attitudes and modes of behaviour, such as good citizenship, co-operation with others, intellectual honesty, scientific attitude etc. Attitudes and interests determine pretty largely the direction of behaviour. Even more than knowledge, attitudes affect action. Many persons even after knowing the evil effects of alcoholic consumption, drink large quantities of alcohol. If however, an emotionally toned attitude is built against it, action follows much more certainly.

Attitude scales are self-report inventories designed to measure the extent to which an individual has favourable or unfavourable feelings towards some person, group, object, institution or idea. They are primarily useful where the individual has little reason for distorting the results, such as in the development of self-understanding or in research.

The most extensive and sustained work in attitude measurement was carried on some 35 years ago by Thurstone and his associates at the University of Chicago. They developed a series of attitude scales each consisting of about 20 statements from extremely favourable to extremely unfavourable. The subject is required to indicate agreement or disagreement with each statement. Topics included in these scales dealt with attitudes toward Negroes, Chinese, war, patriotism, movies etc. Each statement is given a different value according to its importance. An individual's score is simply the median scale value of the statements he has endorsed. A few statements from one of these scales are reproduced below :

Attitude towards Movies

(1) Movies are the most important cause of crime
(2) Movies are the most vital form of art to-day
(3) Movies are good clean entertainment
(4) The influence of movies is decidedly for good

Another method for evaluating pupil's attitudes is the direct observation. When emphasis is laid on the development of attitudes as an instructional outcome e.g., scientific attitude, good citizenship etc., we must rely mainly on observational techniques such as anecdotal records, rating scales and check lists.

III. Projective Techniques

The inventories discussed earlier in this chapter are structured that is, they tend to fix the lines along which a pupil can respond. In contrast to these structured inventories, projective techniques provide almost complete freedom of response. Typically, the individual is presented with a series of ambiguous forms or pictures and asked to describe what he sees. Projective tests are so named because they induce the individual to project himself into the test situation and thus reveal his motives, attitudes, apprehensions and aspirations. He projects his own personality through the responses just as a film projector projects the image on the screen. Hence the name projective tests. There are many projective tests — *sentence completion, doll play, perception of* *ink-blots, interpretation of pictures* and so forth. Despite their differences, all such tests have two features in common.

(1) They present the person with a relatively *non-structured* situation; the situation may be meaningless, ambiguous, amorphous of neutral. The situation does not demand a stereotyped answer as is seen in personality inventories or ability tests. Rigid time limits are also not prescribed. (2) Projective tests have in common the fact that they catch the subject off guard. The purpose of the test is not told to the subject, a frequent practice being to inform him that it is a test of his imagination. The tests are designed to entice the subject into revealing himself without his being aware of the fact that he is doing so. As the subject does not know what is being tested, the test is hard to fake. (3) Most of these tests make little or no demand on literacy or academic skills. They may even be used with children or with illiterate adults. (4) They provide a view of the total functioning individual. They do not slice off one piece or trait for analysis. They preserve the unity and integration of the total personality.

We shall now consider some of the devices under this category :

(*a*) *The Rorschach Test* : This test was developed by the Swiss psychiatrist Hermann Rorschach in 1928. It utilizes ten standard inkblots like the one shown in Fig. 12.2. The subject

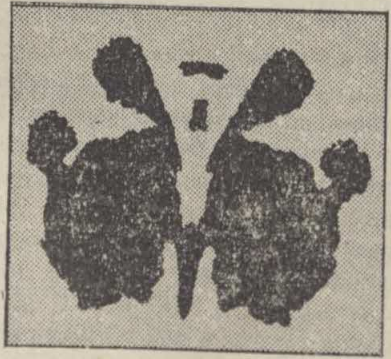


Fig. 12.2 : An inkblot of the type employed in the Rorschach technique.

is shown the ten inkblots, one at a time in a definite order and position. He is asked, "What does it look like? What could this be?" The subject is allowed to turn the blot and look at it from different positions. Different people, of course, *see* different things. Responses are scored in terms of: (i) the total number of items seen, whether the items involve the whole inkblot or only parts, (ii) qualities perceived (colour, form, movement) and (iii) kinds of things reported (like anatomical parts, animals, plants, people and so on).

From the various responses as classified above, the clinician infers certain personality traits. A large number of 'whole' responses may indicate abstract and theoretical traits, whereas a large number of 'small part' responses may indicate a compulsive person occupied with trivialities. If 'movement' responses predominate, the person is thought to have introvertive inclinations, but if 'colour' responses are more numerous, the person is regarded as having warm, free emotional characteristics. All the scoring however, is not quite this simple, for the clinician can interpret not only the number of different kinds of responses but also their pattern.

Scoring of Rorschach data is becoming fairly well standardized, but the interpretation of the scores is far from standard. Most Rorschach testers maintain that the interpreation is necessarily subjective, that it calls for experience, ingenuity, insight and common sense..

(b) *The Thematic Apperception Test (TAT)*: This test was developed by Murray and Morgan in 1938. The material consists of 19 cards containing vague pictures, and one blank card. Like the Rorschach, the TAT is an unstructured experiment; these pictures are ambiguous enough to permit a variety of interpretations by various subjects. The pictures are presented in a definite sequence and the subject is asked to make up a story to fit each picture. The tester says, "Tell me what has led up to the event shown in the picture, what is happening at the moment, what the character in the picture is thinking and feeling and what the outcome will be." In the case of the blank card, the subject is instructed to imagine some picture on the card, describe it and then tell a story about it.

What makes this a useful projective test is that most people when they make up stories, identify themselves with one of the characters in the picture. The subject projects himself into the scene; the responses are dictated by the past experiences, conflicts and wishes of the subject. In this way the subject reveals feelings and desires that he would otherwise hesitate to discuss openly or in some cases would be unwilling to admit to himself.

The TAT unlike the Rorschach, has no standardized scoring system. After the test is over, various themes given by the subject are analysed in an effort to discover the subject's motivational structure. Recurring themes in the stories are noted, e.g., characteristic needs and frustrations of the hero, relations of the hero to members of the opposite sex, to parents or to authority and the overall emotional tone of the stories, whether depressed or overtly optimistic. Thus the interpretation is quite subjective. Content of the stories plays a central role in interpreting the responses. This contrasts with the Rorschach where the centre of attention is not what the subject sees, but how he sees it. Though this test, like the Rorschach uses the projective method, there is a fundamental difference between the two. The Rorschach is intended to reveal the structure and organization (or disorganization) of an individual's personality, while the TAT is devised to bring out primarily the content of one's personality, viz., drives, needs, sentiments, conflicts and fantasies.

A teacher is not expected to use these two devices, since it is an expert's job. A teacher can however use some of the following devices.

(*c*) *Sentence Completion Tests*: This technique has also been used as a projective approach to personality. Any teacher who wants to obtain pupil reactions to the class-room situation and to himself in a relatively unstructured fashion can draw up his own set of incomplete sentence, administer them to his students and apply his own insight in interpreting the responses. The following incomplete sentences might serve the purpose.

1. My teacher............
2. The children in my class............
3. I don't understand..........:..
4. My parents that school............
5. The things that bother me in school............
6. Girls............

Teachers can use these devices to learn what pupils think of parents, teachers, class-mates etc. This will help them to understand the inner feelings of pupils.

(*d*) *Expression through Art* : It is a recognized fact that a child has a natural desire to express himself in some form. The child gives expression to its feeling through such media as chalk, charcoal, coburs, clay, crape-paper etc. The picture drawn by a child may not be meaningful to a layman. But the child's feelings are many times reflected in the picture. Consider the picture of a house drawn by four children. (Fig. 12.3). Are these pictures related in any way to the behaviour traits of these children? The first picture is drawn by Kumud. She does not mix

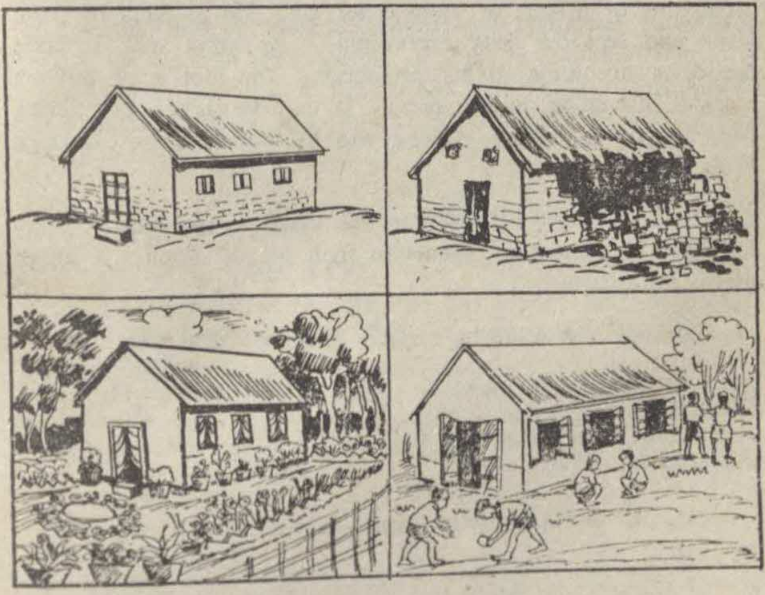


Fig. 12.3 : Expression through Art.

up with her class-mates. She is always silent and tries to keep aloof. The windows and doors of the house are closed. The house is also lonely like its master. The second picture is drawn by Medha who is handicapped because of a fracture in the leg. Her house is in a dilapidated condition. Her feelings of helplessness are reflected in the picture. Karen is very tidy by nature. She likes decoration. Look at the picture drawn by Karen. She has not forgotten to show the window-curtains. Nandini is talkative and quite social. She always likes to play with children. She welcomes her friends to her house. All these characteristics are reflected in the picture.

While studying the pictures drawn by children, we should not look at them from the adult's view point. These drawings supply us a great deal of information about the desires, and repressed feelings of children. It is worth studying how children deface the pictures from their text books Even the scriblings of children are meaningful.

A devotee while worshipping a deity in a temple offers gold ornaments or flowers or some other gift. He projects his own ideas and sees the deity accordingly. An artist will decorate the deity according to his projection. The idol may have a musical instrument in one hand. It may be shown as offering blessings to a devotee. These are all our creations through projection.

(*e*) *Diary* : If children are asked to maintain a diary for recording important events in their life, it supplies a lot of useful information.

Consider the following extract from Nutan's diary.

"All the examples solved by me to-day were correct. My teacher Miss Usha did not appreciate my performance. She praised Sujata's work in the class. I do not like my teacher. She is very unfair in marking and grading."

Here is an extract from Sarita's diary :

"All the girls in my class had put on an attractive dress. I could not get one at home. My mother got irritated when I asked for a new dress. I am an unlucky girl. I must bear my lot."

Such abstracts are so useful that Headmasters and teachers get enough feed back material for knowing their pupils. Teachers can help such children if they can locate the problems.

(*f*) *Other Projective Techniques* : Verbal materials have been used to some extent as media to elicit an individual's projections. In a ' sentence completion test ' the subject is presented with the first part of a number of sentences and is asked to complete them in any way he sees fit, while in a ' story completion test, ' part of a story is read out to the subject, who is then asked to tell what happened in the rest of the tale. Working with clay is a useful projective technique. The characteristic way of dealing with the medium, the vigor of movement and the degree of satisfaction or discontent with the product are all elements that might be involved in the interpretation of behaviour. In the doll play technique the child is provided with a set of dolls representing the various members of a family and is given the materials with which to construct a stage setting. He is encouraged to act out any type of story or scene that appeals to him. What he does and how he treats the playthings is considered to be a reflection of his own personality. In short, a stimulus situation that induces an individual to project the inner aspects of his personality through his interpretations and creations can be used as a projective technique.

IV. Sociometric Devices

A teacher's observations and judgments of pupil behaviour are of special value in some areas, but there are some areas of pupil development where the teacher's evaluation needs to be supplemented by the judgments of pupils, for example, in the field of social adjustment we want to know whether a student is accepted by his fellows and whether they want him around for various activities. In such matters pupils know each other's strengths and weaknesses better than what the teacher knows. The intimate interactions that occur in the *give* and *take* of peer relations are seldom fully visible to an outside observer. For understanding these interactions, peer ratings are often useful. We shall consider two techniques used for peer appraisal. These are also classified as *nominating techniques.*

(a) *The 'Guess Who' Technique*: This technique which is also called as an opinion test, is used to obtain individuals' opinion of one another with regard to a number of traits. The test consists of a series of very brief 'word pictures' each one followed by a blank space in which the child is to write the names of others in the group who may be like the 'word picture.' One may include his own name if he believes the description suits him.

Example

(i) Here is someone who likes to talk a lot, always has some thing to say ....

(ii) Here is someone who always seems rather sad, worried or unhappy, who hardly ever laughs or smiles ....

(iii) Here is someone who waits for somebody else to think of something to do and always likes to follow the suggestions which others make ....

(iv) Here is a girl who likes to read boys' books, play boys' games or would prefer to be a boy ....

(v) Here is someone who can appreciate a joke and see the fun in it even when the joke is on himself or herself ....

(vi) Here is someone who is very friendly, who has lots of friends, who is nice to every body ....

The attractive feature of the nominating pattern is its simplicity which makes it rather painless to administer and usable with children even at the primary stage. An individual's score on a given trait is based on the number of nominations he receives. Students like this nominating technique, if they are assured that no one in the class will see the results. One of the interesting outcomes of this technique is that teachers find differences between their view and the pupils' eye view of the same boy or girl. Teachers see the children from an adult point of view and hence we see a difference in the two views. The teacher can revise her viewpoint and also her approach.

(b) *A Sociogram*: This is a device for studying the social structure of groups. Characteristics of an individual including his popularity may be studied by the *Guess Who* method but the *Sociogram* gives further insight by identifying cliques. hierarchies of leadership and other social groupings. The Sociogram was first developed by Moreno. This technique permits an analysis of each person's position and status within the group, with respect to a particular criterion. A teacher may ask the pupils to name their choices of best friends or of work partners. For example, a teacher might say to a class :

> "We have planned a project in social studies. We need some committees of children who will work together on some part of the project. I would like to know which children you would like to have on a committee with you. Put your name first and under it put 3 or 4 names in order of preference. You are also permitted to list separately any one whom you would not like on the committee."

We now have a series of nominations or choices of work partners. It is possible to show these choices pictorially by a diagram such as that shown in Fig. 12.4. This is called a *sociogram* and the procedure of constructing it is called *sociometry*. The sociogram in Fig. 12.4 reveals several patterns often encountered. There are two groups or cliques. In one Madhuri is the most sought after girl with Kusum, Leela, Suman, Nutan and Shobha as accepted members. In the other group Sujata is the key figure with Kamal, Veena and Sarita as members. Sarita is not thoroughly integrated with the clique; while accepted by Sujata, she is also reaching Madhuri in the other group. Sujata who might be a popular leader of all the girls, instead shows considerable hostility, rejecting three popular girls. Kumud and Neela are fringers, not chosen by any of the others. Rohini and Usha choose each other but are not chosen by any other pupils. Except for the mutual friendship they are isolated. The sociogram thus gives the teacher a framework to understand the behaviour of students. She can modify her views accordingly.

As with the Guess Who method, the teacher must convince the pupils that the sociogram is intended to help them rather than to pry into their affairs. When the teacher finds some children who are isolated in the group, she should try to find out the reasons. The teacher can then help a child become integrated with and accepted by his peer group by (1) providing opportunity for developing friendly relations, (2) improving social skills and (3) building a sense of competence. Teachers have found the sociogram an important method of understanding relations among children so that the school can promote activities which will broaden contacts between children, develop leadership and break down cliques and isolation.

Sociometric patterns are constantly being altered, as friendships waver and regroup. The choices do not indicate any strong and permanent emotional structuring. Similarly the sociogram obtained differs when different problems are asked, for example, if girls are asked to indicate their choice for room mates and their choice of mates with whom to study, the sociometric patterns will differ.

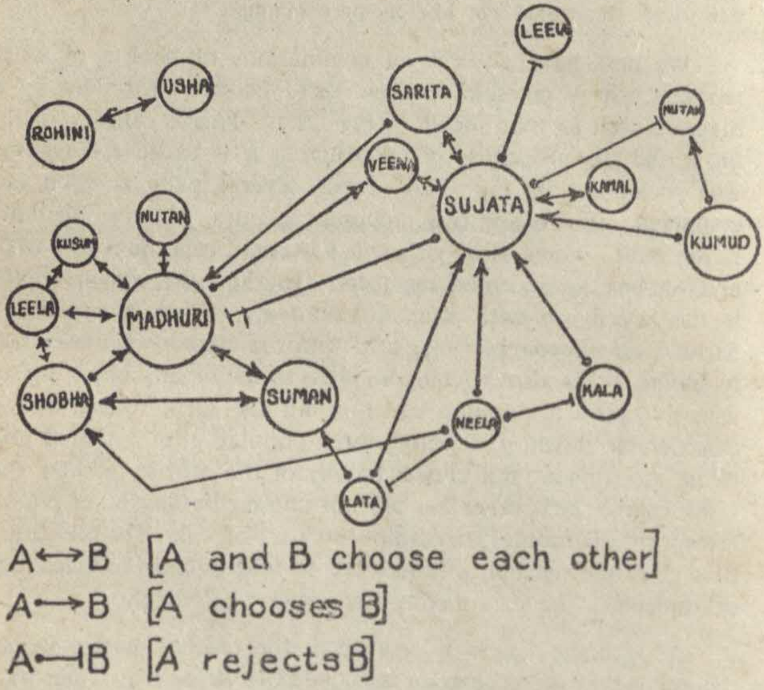


Fig. 12.4 : Sociogram for a class of eighth standard girls.

Concluding Remarks

From the diversity of techniques which have been considered, one impression should be clear. The field of personality testing is still in a formative stage. These instruments have not proved their value empirically to the same extent, as have the aptitude or achievement tests. Peer ratings and self-report inventories provide useful information for understanding pupils better and for guiding their learning, development and adjustment. These purposes will be best served however, when the information is combined with test results, observational data and all other available data concerning the pupils.

It is not expected that a teacher should use all the devices discussed in this chapter. She can often use the check list and the rating scale, so that qualitative descriptions of pupil behaviour are made available for evaluation purposes.

FOR FURTHER STUDY

1. Explain the various types of rating scales. What are the common errors in rating ? What are the requirements of a good rating scale ?
2. What are the various observational techniques used to evaluate personality traits ? Discuss the merits and limitations of each.
3. Prepare a descriptive graphic rating scale for rating *sociability and emotional control* of the students in your class. What precautions will you take to minimize the errors of rating ?
4. Discuss the limitations of :
   - (i) Personality inventories,
   - (ii) Interviews
   - (iii) Anecdotal records.
5. What does the name projective test signify ? How does it differ from an inventory ? Discuss the common features of the various projective tests.
6. Illustrate a sociogram. What is its importance to a teacher ?
7. How can a teacher make use of the nominating techniques for understanding children ? What precautions should be taken while using these methods ?
8. What is the difference between a 'check list' and a 'rating scale' ? What important considerations should be borne in mind while devising and using them ?
9. Explain the importance of an anecdotal record. What precautions should be taken in writing such a record ? Discuss the advantages and limitations of anecdotal records.
10. What is the difference between a personality inventory and an interest inventory ? What are their merits and limitations ?
11. Announce the following problem to your class and prepare a sociogram :

"We want to form study groups within the class for mathematics. Indicate four choices with whom you would like to study and also four names with whom you would not like to study."
How does such a sociogram help the teacher in modifying her views about the students?

12. Explain the difference between:
    (i) a rating scale and a personality inventory.
    (ii) 'guess who' technique and a sociogram.
    (iii) a rating scale and a check list.
13. Which of the following traits would probably be hardest to rate reliably after observation? Which one the easiest? Why?
    (i) talkativeness (ii) freedom from tension (iii) freedom from anxiety (iv) leadership (v) sensitivity to criticism (vi) teaching ability.
14. Prepare a check list to study the following characteristics:
    (i) Laboratory skill of students
    (ii) Study habits of children
    (iii) Teaching ability of student-teachers
    (iv) Performance of candidates in elocution competition.

* * *

13

INTERNAL ASSESSMENT

We have seen in chapter ten that the public examination dominates the school programme tremendously. As a result of this domination, there is a feeling among the students that the outcome of the annual examination alone will push them forward or will pull them down. All sorts of corrupt practices, mass copying, threats to invigilators, bribes and what not, are likely to prevail. Honesty is at a discount and fraudulent practices at a premium.

To remedy this situation, Educationists have made three suggestions, namely improving the question papers, improving the procedure of scoring and introducing the concept of internal assessment. We have discussed the first two suggestions in chapter ten.

Importance of Internal Assessment

Various programmes are carried out during an academic year. We believe that through these programmes there will be a change in behaviour in respect of knowledge, skills, attitudes, interest etc. Promotion of a student from one class to another is at present mostly determined on the basis of the annual examination Pupils therefore can afford to neglect the day-to-day activities, daily assignments, group projects etc. A pupil gets through the examination with utter disregard to many of the desired qualities such as development of proper study habits, interest in reading, laboratory skills, and personality traits such as co-operation, initiative, honesty, confidence, hard work, perseverance etc. By introducing the system of internal assessment, we can change the attitude of the students favourably toward the day-to-day activities of the school.

Areas of Internal Assessment : Internal assessment should help a teacher to answer such questions as :

1. Is Geeta progressing well in various school subjects ?
2. Does she participate in co-curricular activities such as debating, dramatics, sports, excursions and other group activities ?
3. Is she developing qualities such as co-operation, initiative, perseverance, consideration for others etc. ?
4. Is she regular in completing daily assignments ?
5. Does Geeta read books from the library ? Has she developed interest in reading ?
6. Has she developed proper study habits ?
7. Has she developed laboratory skills ?
8. Can she recite poems ?

It will thus be seen that internal assessment can cover a number of areas of pupil's development. For promotion purposes schools must take into consideration the performance of students in various areas mentioned above. We have to assign some weightage to various activities that will cover the above areas.

Plan of Internal Assessment

Having decided the areas we can prepare a tentative plan of operation.

(i) *Periodical Tests* : These tests should be given to see the achievement of pupils in various academic subjects. The subject teacher should give the test during one of her teaching periods so that the school routine will not be disturbed. At least four tests should be given in each subject in a year, each test carrying a maximum raw score of twenty five.

(ii) *Study Habits* : Home work is generally given in all subjects. Some weightage should be given to the performance of students, who do this work regularly. About ten assignments in each subject should be checked, each being given a maximum score of ten.

(iii) *Participation in Sports* : A student should participate at least in three items from athletics. Those who participate successfully in more than three items should obviously be given more credit than the minimum.

(iv) *Oral Work* : Recitation of prose passages, poems etc., should be made obligatory. Performance should be assessed at least once in each term.

(v) *Laboratory Work* : Qualities such as accuracy, neatness, regularity etc., should be observed in the laboratory work. One test may be held in each term.

(vi) *Co-curricular Activities* : A student should participate at least in two activities in any of the special fields namely music, drawing, elocution, mimicry, dramatics, story-telling, school magazine, exhibition of models and crafts and any other projects undertaken by the class.

(vii) *Library Work* : A student is expected to read books and develop interest in reading. He should maintain a record of what he reads. He should write in about ten lines the description of every book that he reads. Every student should read at least two books in a term.

(viii) *Personality Traits* : Traits such as co-operation, initiative, honesty, leadership, perseverance, confidence etc. should be observed and assessed.

Weightage

Weightage to the different areas may be given as shown below : Scores in the periodical tests of a subject in a year should be added and finally reduced to a total of 25 scores. These should then be added to the marks obtained in the annual examination.

Final score in a subject = Score in periodical tests + Score in the final examination
(maximum 100) (maximum 25) + (maximum 75)

Scores in various other items should be added as shown below. A student must obtain a minimum score of 50 in order to get through.

1. Study habits	10
2. Participation in sports	20
3. Oral work	10
4. Laboratory work	10
5. Co-curricular activities	20
6. Library work	10
7. Personality traits	20
Total	100

The above scheme will help reduce the undue emphasis given to the final examination in academic subjects. Periodical tests will make the students alert. Students will also pay attention to various other activities since a minimum score of 50, is expected of every student in these activities. Performance of students in various areas mentioned above should be recorded in the progress card (Fig. 9.2, page 175). For assessing the achievement of students in various areas, teachers should use observational techniques such as the rating scale or the check list described in chapter XII.

Validity of Internal Assessment

Internal assessment is a powerful tool in the hands of a teacher. It is likely to be misused. There is every possibility that a teacher may be tempted to rate a favourite student higher than another in various areas, simply because the student is her favourite. We have to guard against our own prejudices.

Gloria gets a score of 76 while Rose gets a score of 47 in internal assessment done by their teacher Miss Margaret. On careful analysis it was revealed that Gloria wishes her teacher *good morning* every day. She is very submissive and does whatever is told by her teacher. Miss Margaret is influenced by Gloria's attitude. Rose is equally a good student but is not so submissive as Gloria. Rose is also not very particular in wishing her teacher a *good morning*. She however is not arrogant. The teacher however is prejudiced against Rose. Teacher's assessment of these two girls is not valid. She has assessed some qualities other than those mentioned in the plan of internal assessment. Hence the validity is at stake. We fall a prey to our prejudices. The evaluator must be very careful while assessing these qualities. She should be as objective as possible in assessing the performance. If students find that the teacher is unduly influenced by such triflings, they will pay attention to this aspect and will use foul means to please the teacher. Students will consider *pleasing the teacher* as an important objective of the school programme.

Internal Assessment in Training Colleges

The practice of internal assessment is usually followed in all training colleges. Marks are assigned to such items as practice teaching, preparation of teaching aids, case study, participation in co-curricular activities etc. Maximum weightage is given to the programme of practice-teaching. Lessons given by the trainees are observed by the method-masters. Subjectivity is likely to creep in if method-masters are not given a check-list for assessing the performance of trainees in teaching. A sample proforma for lesson-evaluation is given below. The method-master should write the remarks in the blank space provided.

Lesson Evaluation Form

Points for Observation	Remarks of the Method-Master
I. Lesson Planning	
1. Is the lesson well planned ?	
2. What are the significant teaching points emphasized ?	
3. Is the sequence logical ?	
4. Are pupil-activities emphasized ?	
II. Introduction	
1. How quickly does the teacher take control of the class ?	
2. How is the lesson introduced ?	
3. Are the pupils motivated effectively ?	
III. Teaching	
1. Is the teacher clear about her objectives ?	
2. What concepts, understandings and generalizations are developed ?	
3. What skills are emphasized ?	
4. Are the questions thought-provoking	
5. What questions and problems do pupils raise ?	
6. Did pupils participate sufficiently ?	
7. What provision is made for handling individual differences ?	
8. Is there clarity of thought in the presentation ?	

Points for Observation	Remarks of the Method-Master
9. Are pupils given an opportunity to show initiative ?	
10. What steps are taken to locate individual difficulties and how are they dealt with ?	
11. Are visual aids used effectively ?	
12. Are the assignments appropriate ?	
13. How is the lesson summarized ?	
14. How are the facts, understandings and skills tested ?	
IV. Teacher's Personality	
1. Are speech and manners effective ?	
2. Are the language and the style appropriate to the level of students ?	
3. Does the teacher control the class effectively ?	
4. Is she resourceful, pleasant and tactful ?	
5. Does she show an understanding of pupil's behaviour ?	
6. What kind of pupil-teacher relationship exists ?	
V. Summary	
1. What are the strong points of the lesson ?	
2. What are the weak points ?	
3. What suggestions would you offer ?	
VI Final Evaluation	
How successful is the lesson in your opinion ?	
Use the following ratings :	
A : very effective B : very good	
C : average D : needs much improvement.	
E : very poor.	
	Signature of the Method-Master

The above proforma is a lengthy one. We can make it more compact by using a rating scale as shown below.

Rating of the student-teacher's practice lesson.

	A 4	B 3	C 2	D 1	E 0
1. Method of introduction					
2. Development of concepts, understandings and skills					
3. Pupil's involvement in the learning process					
4. Interest created					
5. Provision for individual differences					
6. Clarity of thought					
7. Class management					
8. Mastery over the subject matter					
9. Blackboard work					
10. Use of teaching aids					
11. Total effect of the lesson	A 10	B 8	C 6	D 4	E 2

Suggestions :

Total Score :

Signature :

For each of the items 1 to 10 of the above scale the scores range from 0 to 4, while the scores for the last item range from 2 to 10. Maximum score will be 50 and minimum will be a score of two.

The rating scale will help the method-master to assess the lessons objectively. We can use such rating scales for evaluating other activities of the student-teachers. About 300 marks are allotted to the year's work done by the trainees. We have to be very watchful in using this instrument of internal assessment. There have been examples where marks have been given indiscriminately and indiscreetely. In such a case the marks lose their sancity. No wonder that the internal assessment done in this way creates a distrust among the students and they use short-cut methods to boost up their result.

Guard against over-enthusiasm : Sometimes an over-enthusiastic head of an institution prepares a very exhaustive check list or a rating scale for bringing in objectivity in the assessment. For rating the practice-lesson we have given above a rating scale consisting of eleven items. For making it more objective, more items can be added and the scale can be made more comprehensive by including in it about 50 to 60 items. Every minute detail of teaching can be included in the list Such a scale will help attain objectivity but it will lose its usability (Chapter IX). An instrument must be usable.

Advantages of Internal Assessment

1. Teachers will not give undue weight to external examinations.
2. Students will pay attention to daily assignments and to all activities organized in the school.
3. Proper study habits are likely to be developed.
4. A comprehensive picture of pupil's progress can be obtained.
5. It enables teachers to diagnose pupils' weak areas.
6. It brings out a change in the attitude of students and teachers toward the school programme.

Limitations of Internal Assessment

External examinations serve as a motivating factor in the life of students. Internal assessment should therefore supplement the external examination and should not try to replace it. Internal assessment is like a sharp instrument. It can cause great harm in the hands of an inefficient technician. If not properly handled, it will do injustice rather than justice to the students. When favouritism and personal prejudices are rampant, internal assessment will not be valid. A favourite student is likely to get a high score than the one, not liked by the teacher. Students will then raise the slogan " Down with Internal Assessment. " It is difficult to make this instrument completely valid, reliable and objective as described in Chapter IX. However we can develop these qualities by adopting an analytic approach while observing students' behaviour.

Retrospect and Prospect

We have discussed in the preceeding chapters the meaning of evaluation and its various techniques. Educational evaluation plays an important role in the school. It is an integral part of the instructional programme and it provides information which serves as a basis for a variety of educational decisions. The main emphasis in evaluation however, is the pupil and his progress in learning.

Evaluation does not merely aim at improving the examination. That would be a very limited function. It should improve the teaching-learning process. A teacher must always take into consideration the interrelated nature of teaching, learning and evaluation. This relationship can be seen if we study the sequential steps in the process of education mentioned below :

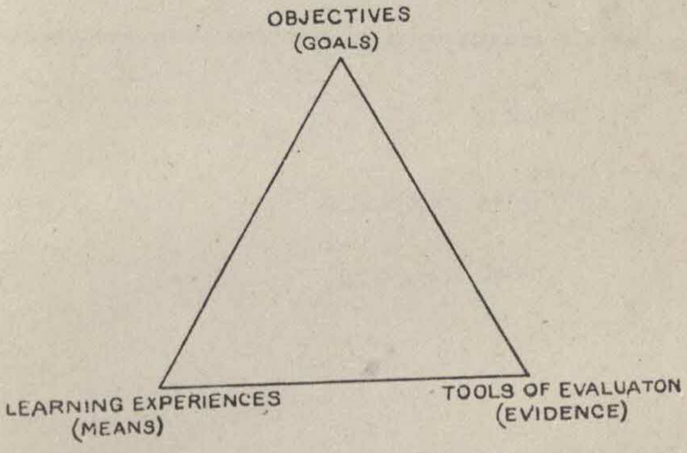


Fig. 13.1 : Triangle of Education.

(i) identifying and defining objectives
(ii) planning and directing learning experiences
(iii) determining pupil progress
(iv) using the results to improve learning and instruction.

Educational evaluation should be viewed as a process and not merely a collection of techniques. If the teacher accepts this view and adopts right type of attitude, she will be able to accept the challenge and improve the teaching-learning process.

FOR FURTHER STUDY

1. Discuss the importance of internal assessment.
2. Prepare a plan of internal assessment for pupils of standard VIII.
3. You are asked to assess the pupils' work in the following areas : (i) library reading, (ii) development of personality traits such as initiative and attitude of co-operation. Discuss how you will assess these areas.
4. Discuss the factors that affect the reliability and validity of internal assessment.
5. Discuss the merits and limitations of the system of internal assessment.

* * *

Section II

Statistical Methods

14. Frequency Distributions

15 Measures of Central Tendency

16 Measures of Variability

17 Measures of Relationship

18 The Normal Curve

14

FREQUENCY DISTRIBUTIONS

Progress in science often depends upon the developments of quantitative methods. Without such methods science is limited to crude observation and classification. With them, it can greatly extend and refine the conclusions it can draw from its data. Although the problems of measurement are much the same in all sciences, psychologists are probably more concerned about the logic of measurement than are most other scientists. The reason for this is that many of the things they want to measure are rather complex. Courage, for example, is not the same sort of thing as the length of a table; there is no simple yardstick you can place beside a man to measure his courage. Psychologists have had to devise roundabout ways of getting at what they want to measure. They have devised a number scales for measuring human characteristics.

Once we have a scale for measuring some human characteristic, our next problem is to know what to say about the measurements we make. This is an especially difficult problem because people differ. They differ in height, weight, intelligence, learning ability, attractiveness and a number of other traits, and the scores representing objective measures of these traits vary widely. Any attempt to describe, summarize or compare results for individuals or for groups calls for numerical treatment. This branch of mathematics that deals with the analysis of sets of scores for groups of individuals is known as 'statistics.' In the following chapters, we shall describe some of the statistics that we use in interpreting psychological and educational measurements.

How can statistics be useful to a teacher?

Suppose, you have given tests in science and mathematics to two classes of eighth grade. You have scored the papers and recorded the results. The data should provide answers to your querries about these classes. Let us consider a few of the questions that can be answered with the help of these data.

(1) What is the general pattern of the set of scores? How do they run? To answer this question we need to consider simple ways of tabulating and graphing a set of scores.

(2) A second type of question that would arise is: What is this group like, on the average? Have they done as well on the test as other eighth grade groups? What is the typical level of performance in the group? All these questions call for some single score to represent the group as a whole or some measure of the middle of the group.

(3) Further you might feel a need to describe the extent to which the scores spread out away from the average value. Are all the children in the group about the same, so that the same materials and procedures would be suitable for all? If not, how widely do they spread out on a given test? How does the group compare with other classes with respect to spread of scores?

(4) Next you may be interested to know about the standing of a particular student. Has Mangesh done well or poorly in science? How good or poor is his score? Has he done better in mathematics or in science? To answer this question we need a common yardstick in terms of which to express performance in two quite different areas.

(5) A fifth question may be of this type: To what extent do those who excel in science also excel in mathematics? To what extent do the abilities in these two subjects go together in the same individuals? If Nandini is superior in one, is she likely to be superior in the other subject? The measures of this *going-togetherness*, we speak of as measures of correlation.

(6) From the distribution of scores, we can also say something about the test. How are the scores distributed? Are they piling up at the low end or at the upper end? Are the questions too difficult or too easy? Or do they vary uniformly in difficulty?

This following chapters will be devoted to illustrate the methods developed for answering these questions. There are many more questions that arise with respect to such data. A group of 80 boys surpasses a group of 80 girls from the same school on a mathematics test. This is a descriptive fact true of these particular groups. We would like to know whether we can safely conclude that all boys of that age group would surpass all girls of the same age group on the above test. This is a problem of inference. Problems of statistical inference make up the bulk of advanced statistical work, but we shall be mainly concerned with descriptive statistics in this volume.

Misuse of Statistics : Every one has heard such gibes as ' There are little liars, big liars and statisticians, ' " You can prove anything provided you use statistics. " and ' Statistics supports many mistaken things including statisticians. ' It must be admitted that statistics in the hands of a beginner is an uncertain tool and in the hands of a propagandist, a dangerous one. Many examples of numerical data admit of more than one kind of analysis and if a person sets out to analyse any given evidence according to some point he wishes to establish, he probably will be able to construct an argument plausible enough to fool at least the ignorant. Following are some of the illustrations : (i) Using an average value to represent a set of numerical data when the average obscures the important features of the data, e.g., if the monthly salaries of three workers are Rs. 700, Rs. 105 and Rs. 95 respectively, the arithmetic mean of Rs. 300 would obscure more information than it would convey. If the mean were the only information reported, the data would in effect be falsified. (ii) Making comparisons without reference to all the pertinent data : To say that city-teachers are better paid than rural teachers has little meaning unless cost of living is taken into account.

Throughout the following chapters considerable attention will be given to the proper use of statistics.

Classification and Tabulation of Test Scores

1. *Meaning of Test Scores* : In measuring mental and physical traits most of the variables with which we deal, fall into continuous series. A continuous series is one which is capable of any degree of subdivision, though in practice divisions smaller than some convenient unit are rarely met. Thus it is possible to have scores like 74·7 or 74·8 or 74·9 on various tests, though usually they are thought of as increasing by increments of one unit such as 74, 75, 76 etc. Scores in a continuous series are to be

thought of as distances along a continuum rather than as discrete points. A score of 74 represents the interval 73·5 to 74·5. The exact midpoint of this interval is 74. Other scores may be interpreted in the same way. A score of 50, for instance, includes all values from 49·5 to 50·5. This means that 49·8, 50·0 and 50·4 would all be scored 50. The usual mathematical meaning of a score is an interval which extends along some dimension from ·5 unit below to ·5 unit above the face value of the score.

2. *Grouping of Scores* : An intelligence test is given to the students of a primary school. The record of scores is reproduced in Table 14.1. Let us look at the scores and consider how they could be rearranged so as to give us a clear picture of how the pupils did on this test.

Table 14.1 : Scores on an Intelligence Test.

91	60	80	100	71	120	95	91	119	106
101	62	97	93	89	126	73	129	90	84
81	114	117	64	88	98	92	108	109	75
130	94	84	116	138	98	108	81	101	99
114	77	124	95	91	116	99	82	107	72
124	83	97	95	115	108	87	102	75	128
82	114	86	106	77	94	96	86	123	114
83	103	84	92	112	105	88	111	104	93

The simplest rearrangement would be just to arrange them in order from highest to lowest. We may then have something that looked like this :

138, 130, 129, 128, 126, 124, 124, 120, ...
.. .. 73, 72, 71, 64, 62, 60

This arrangement gives a somewhat better picture of the way the scores fall. We can see the highest and the lowest scores at a glance. But this simple arrangement of scores still has too much detail for us to see the general pattern clearly. It is also not a convenient form to use in computing. We need to condense it into a more compact form.

Table 14.2 : Grouping of Test Scores.

Interval of 1 unit (a)		Interval of 5 units (b)	Interval of 10 units (c)	Interval of 20 units (d)
Scores f	Scores f	Scores f	Scores f	Scores f
138...1	122....	135–139... 1	130–139... 2	120–139... 9
137...0	121....	130–134... 1		
136...0	120...1	125–129... 3	120–129... 7	100–119...25
136...0	:	120–124... 4		
135...0	: :	115–119... 5	110–119...11	80– 99...36
133....	75...	110–114... 6	100–109...14	60– 79...10
132....	74...0	105–109... 9	90– 99...20	
131....	73...1	100–104... 5	80– 89...16	
130...1	72...1	95– 99...10	70– 79... 7	
129...1	71...1	90– 94...10	60– 69... 3	
128...1	: :	85– 89... 6		
127...	: :	80– 84...10		
126...1	62...1	75– 79... 4		
125....	61...0	70– 74... 3		
124...2	60...1	65– 69... 0		
123...1		60– 64... 3		

A further step in organizing the scores for presentation is to prepare what is termed a *frequency distribution*. Four different ways in which the same eighty scores can be classified into a frequency distribution are shown in Table 14.2. The first illustration 'a' in which the scores retain their individual identities may be called a simple or ungrouped frequency distribution. It shows how often each score occurred. Only a portion of the frequency distribution is shown; the table will be too long and spread out. This is not a convenient form of distribution. The remaining three illustrations b, c and d in which the grouping of scores destroys the individual identities of most of them, are called *grouped frequency distributions*. When individual scores are not important a statistician makes them more compact by squeezing them into classes. This grouping of scores into classes is called *a frequency distribution* because it permits us to count the frequency of scores in each of the classes and then to see how the frequencies are distributed among the various classes.

Out of the four illustrations in Table 14.2, illustration 'a' is not suitable, as the distribution cannot be easily obtained. Illustration 'd' furnishes the basis for obtaining quick but quite unsatisfactory information, for the very rough grouping almost entirely sacrifices even the approximate identity of the individual scores. Illustrations b and c, neither of which demands an undue time expenditure in order to obtain accuracy nor sacrifices accuracy for a saving in time and labour represent satisfactory practices midway between the two extreme methods of handling the scores. In the preparation of a frequency distribution, the method of grouping test scores is somewhat similar to that followed by postal clerks in the distribution of outgoing letters in a large post office. Mail designated for certain sections of the country is thrown into the proper mail pouches. Number of pouches required depends on the number of pieces of mail to be distributed and also on the population that can be effectively served by a given pouch.

Steps in preparing a Frequency Distribution

(a) *Determine the range* : Find the highest score and the lowest score in the series. Find the difference between these scores. The difference is called the range (R) of the series. In the case of scores given in Table 14.1, the range is 138–60=78. The range is useful in determining the number of class intervals to use in the frequency table.

(b) *Determine the size of the class-intervals* : Use the range to determine whether the scores should be classified by units of 1, 3, 5, 8, 10 or 15 i.e., to determine the size of the class intervals. No special rule can be stated. The decision is a compromise between losing detail from our data on the one hand and obtaining a convenient, compact and smooth representation of our results on the other. A broader interval loses more details but condenses the data into a more compact picture. This will be evident from illustration 'd' in Table 14.2. An interval of 3 units will yield about 27 classes in our example, an interval of 5 would give 16 classes, while an interval of 10 would give 8 classes. An interval of 3 would spread the data out too much, thus losing the benefit of grouping, where as an interval of 20 would crowd the scores into too coarse categories. Accordingly an interval of 10 is chosen as best suited to the data of Table 14.1.

(c) *Tabulate the Scores* : It is not necessary to take zero as the starting point. In the above example we can conveniently start from 50. If the lowest score were 44, we could start preparing the table from 40 onwards. It is an accepted custom to place the highest scores at the top of the list and the lowest at the bottom.

Table 14.3 : Scores from Tables 14.1, grouped into a frequency table.

(1) Class-Intervals	(2) Tabulation	(2) Frequency (f)
140–149	—	0
130–139	//	2
120–129	𝍸,//	7
110–119	𝍸,𝍸,/	11
100–109	𝍸,𝍸,////	14
90– 99	𝍸,𝍸,𝍸,𝍸	20
80– 89	𝍸,𝍸,𝍸,/	16
70– 79	𝍸,//	7
60– 69	///	3
50– 59	—	0
		N = 80

Having decided upon the size of the class interval and the starting point, we are ready to list the scores as in Table 14.3. Begin with the first score in the original list of scores. Place a tally mark in the tabulation column (Table 14.3) of the appropriate class interval. The first score of 91 is represented by a tally placed opposite interval 90–99, the second score of 60 by a tally placed opposite interval 60–69. The remaining scores have been tabulated in the same way. In order to facilitate counting, every fifth mark is made slanting across the preceeding four tally marks, as shown in Table 14.3. When all the eighty scores have been listed, the total number of tallies on each class interval is written in column 3, headed f (frequency). Frequency 'f' denotes the number of cases in a class interval. It appears from Table 14.3 that there are 20 students who received scores between 90 and 99. The number 20 is therefore the frequency for the class interval 90–99. It should be obvious that the sum of all the class-frequencies is equal to the total number of cases in the table. The sum of the 'f' column is called N. When the total frequency in each class interval has been tabulated opposite the proper interval as shown in column 3, our eighty scores are arranged in a frequency distribution.

Advantages and Limitations

In making such a frequency distribution, something is lost and something is gained. What is lost, is the identity of individual measurements; we cannot, for example, any longer tell from Table 14.3, what the score is for the twenty-eighth man. We also do not know the precise value of any individual measurements. We cannot tell, for example, what the sixteen measurements were like in the interval 80–89. We do not know how many of them are 80's, how many 81's, how many 82's and so on We have lost this information in the grouping. However, the losses are more than offset by what is gained. The most immediate gain is an increase in the information conveyed by the frequency distribution about the group as a whole. We can see at a glance that the largest group has scores between 90 and 99, while the majority of the students have scored between 80 and 109. Thus the construction of the frequency distribution has helped us to appreciate quickly the general nature of our measurements. If Manda has a score of 109, we can estimate her rank and also can express a qualitative judgment about the rank.

It is now possible to answer by mere inspection of the table. such questions as:

(i) How many students obtained scores between 100 and 109? ...... (14)

(ii) How many obtained scores above 99? ...... (34)

(iii) What per cent of the whole group obtained scores between 100 and 109? ...... (17·5%)

(iv) How many students have their scores between 60 and 79? ...... (10)

(v) Is a score of 95 high, average or low? ...... (average).

Limits of Class Intervals

There are various ways of expressing the limits of class intervals in a frequency distribution. Table 14.4 illustrates three different ways.

Table 14.4 — Methods to describe the limits of Class Intervals.

A		B		C	
Class Intervals	f	Class Intervals	f	Class Intervals	f
120–130	7	120–129	7	119·5-129·5	7
110–120	11	110–119	11	109.5–119.5	11
100–110	14	100–109	14	99.5-109.5	14
90–100	20	90– 99	20	89.5– 99.5	20
80– 90	16	80– 89	16	79.5– 89.5	16

In (A), the interval 80 up to 90 means, as we have already seen that all scores from 80 up to but not including 90 fall within this grouping. The intervals in (C) cover the same distances as in (A) but the upper and lower limits of each interval are defined more exactly. We have already seen that a score of 80 on a continuous series ordinarily means the interval 79·5 up to 80·5 and that a score of 89 means 88·5 up to 89·5. Accordingly to express precisely the fact that an interval begins with 80 and ends with 89, we may write 79·5 as the lower limit and 89·5 as the upper limit of this interval. The class intervals in (B) express the same facts more clearly than in (A) and less exactly than in (C). Thus '80-89' means that this interval begins with score 80 and ends with score 89, but the precise limits of the interval are not given. Remember that in all the three methods, ten scores are included in each interval. Method (B) is the easiest for tabulation, but one must remember that the *expressed* class-limits are not the *actual* class-limits; thus interval 80–89 begins at 79·5 (not 80) and ends at 89·5 (not 89). Hence the size of the interval is 10 and not 9. The same principle holds no matter what the size of the interval or where it begins. An interval labelled '14–16' includes scores 14, 15 and 16 and extends exactly from 13·5 to 16·5. An interval labelled '71–72' extends from 70·5 to 72·5. Method B will be followed throughout this book. For finding out the size of the interval, always consider the precise limits of the class-interval as shown in C.

The mid-point of an interval : The scores grouped within a given interval in a frequency distribution are considered to be spread evenly over the entire interval. If we wish to represent all of the scores within a given interval by some single value, the mid-point of the interval is the logical choice. A simple rule to find out the mid-point is to average the expressed or the actual class-limits. The interval containing scores 80 to 89 inclusive has precise limits of 79·5 and 89·5. Hence the mid-point of this class interval is $\frac{80+89}{2}$ or $\frac{79\cdot5+89\cdot5}{2} = 84\cdot5$. Mid-point of the class interval 14–16 would be 15 while that of the interval 70–74 would be 72. The assumption that the midpoint in the most representative value within an interval holds best when the number of scores in the distribution is large and when the intervals are not too broad. But even when neither of these conditions fully obtains, the mid-point assumption is not greatly in error and is the best we can make. In the long run, about as many scores will fall above as below the various midpoint-values and lack of balance in one interval will usually be offset by the opposite consideration in another interval.

Some authors follow procedure A as given in Table 14.4. In that case an interval of 80–90 includes all scores from 80 up to 90 but not the score 90. The midpoint of this interval would then be

85. Thus the real limit of the class interval and its midpoint are shifted by ·5 of a score point upward, if method A is followed.

Graphic Representation of Frequency Distributions

In an earlier section, it was pointed out that one can extract information from a frequency table more readily than from a heterogeneous list of scores. Aid in analysing numerical data may often be obtained from a graphic or pictorial treatment of the frequency distribution. The advertiser has since long used graphic methods because these devices catch the eye and hold the attention which the most careful array of statistical evidence fails to attract notice. For this and other reasons the research worker also utilizes the attention getting power of visual presentation and at the same time, seeks to translate numerical facts — often abstract and difficult of interpretation — into more concrete and understandable form. Thus the advantages of graphical representation are two fold, first it is a pictorial vivid representation of facts and secondly it helps to dilute the abstractness of ideas.

We shall consider only two methods of representing a frequency distribution graphically namely, the *histogram* and the *frequency polygon.*

(a) *The Histogram or Column Diagram* : The purpose of a histogram is to show the frequencies within classes graphically. In a histogram, it is assumed that scores are spread uniformly over their intervals. The frequencies within each interval of a histogram are represented by a rectangle, the base of which equals the size of the interval and the height of which equals the number of scores (f), within that interval. Thus the three scores within the interval 60–69 (Table 14.3) are represented by a rectangle whose base equals the length of the interval and whose height equals three units measured off on the Y axis (Fig. 14·1). The seven scores within the next interval 70–79 are represented by a rectangle one interval long and seven Y units high. Thus at every class interval we draw a column, the height of which is proportional to the frequencies. The highest rectangle, of course, is on the interval 90–99 which has 20, the largest frequency, as its height. Such a graph is also known as a *column graph.*

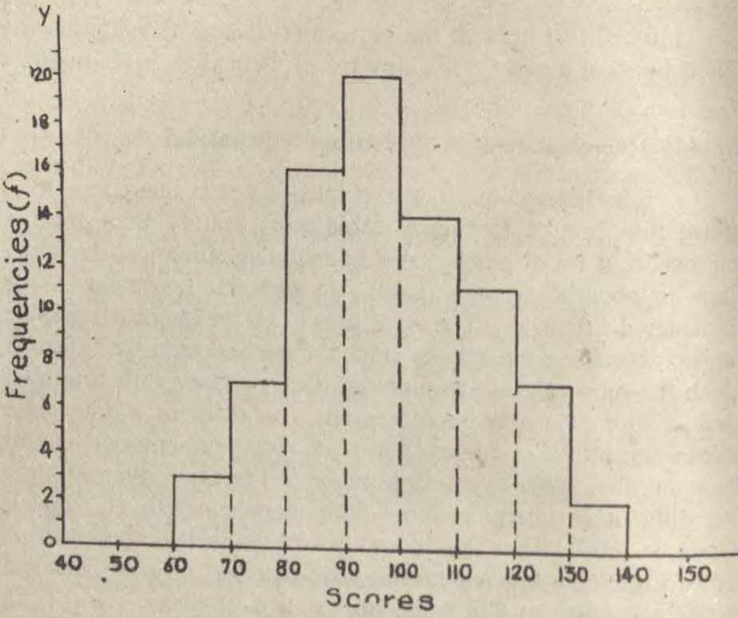


Fig. 14.1 : Histogram of Scores given in Tables 14.1 and 14.3.

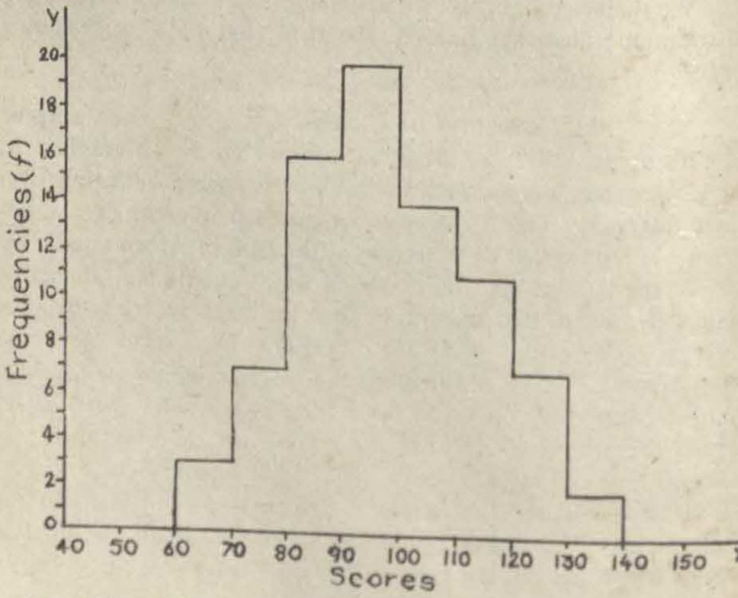


Fig. 14.2 : Histogram of Scores given in Tables 14.1 and 14.3.

In Fig. 14.1, the precise class-limits are not shown, e.g., the interval 60 to 69 extends from 59·5 to 69·5; a rectangle should be erected on this interval. But for convenience, the integral numbers 60, 70, 80, 90 etc. have been used; they correspond to the exact limits of 59·5, 69·5, 79·5 and so on. While each interval in a histogram is represented by a separate rectangle, it is not necessary to project the sides of the rectangles down to the base line as is done in Fig. 14.1. Only the boundary line as shown in Fig. 14.2 may be retained. The rise and fall of the boundary line shows the increase or decrease in the number of scores from interval to interval and is usually the important fact to be brought out.

The histogram has one interesting property, namely, that the area of a histogram corresponds to the total frequency (N) of the distribution and that the area of each rectangle in a histogram corresponds to the frequency within a given interval. Verify this property by comparing the frequencies in Table 14.3 with the area of each rectangle in Fig. 14.1.

(b) *The Frequency Polygon* : If the mid-points of the upper bases of the rectangles in a histogram are connected, the resulting figure is *a frequency polygon* and is illustrated in Fig. 14.3.

A histogram and a frequency polygon differ in one fundamental point. In a histogram, we have assumed that the scores are spread uniformly over their intervals, while in a frequency polygon we assume that all scores in a given interval are at the mid-point of the interval. Take for instance the class interval 110 to 119. There are 11 students who have scores within the range of 110 to 119. For practical purposes we assume that all the 11 students have a score of 114·5 which is the mid-point of the class interval 110–119.

Fig. 14.3 illustrates the method of plotting the points. Mid-points of class-intervals are plotted on the X-axis, and the frequencies (*f ' s*) are plotted on the Y-axis. Frequencies on each interval are plotted above the midpoints of the intervals on the X-axis. Frequency of the class interval 60–69 is 3. To represent this score on the graph, we go out on the X-axis to 64·5, midpoint of the interval and count up 3 units along the Y-axis and put a dot there. Frequency on the next interval 70–79 is 7, so the

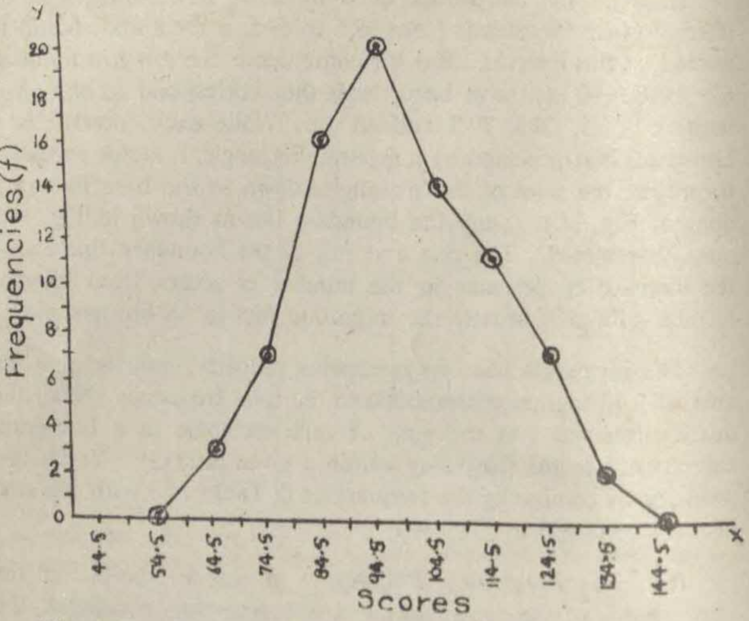


Fig. 14.3 : Frequency Polygon of Scores given in Table 14.3.

second point must be plotted above 74·5 and 7 units up on the Y-axis. In this way, in plotting a frequency polygon the midpoint of an interval is always taken to represent the entire interval.

When all the points have been located, they are joined in regular order to form the frequency polygon as shown in Fig. 14.3. In order to close the polygon that is, to bring it down to the base line, one additional interval (50–59) at the low end and one additional interval (140–149) at the high end of the distribution are included on the X-axis. The frequency on each of these intervals is zero; hence by adding them to the scale, we begin the polygon one half interval below the first and end it one half interval above the last interval on the X-axis.

How does a frequency polygon differ from a histogram ? :

(i) In a histogram, it is assumed that the scores are spread uniformly over their interval while in a frequency polygon we assume that all scores in a given interval are at the mid-point of the interval. Thus the eleven students in the interval 110–119 are assumed to have a score of 114·5 which is the mid-point of 110–119. (ii) The total frequency (N) of a distribution is represented by the area of its polygon, i.e., the area bounded by the frequency surface and the X-axis. That part of the area lying above any given interval, however, cannot be taken as proportional to the number of cases within the interval owing to irregularities in the frequency surface. In this respect, it differs from the histogram, which presents an accurate picture of the relative proportions of the total frequency from interval to interval.

Which graph will you prefer? : As is to be expected, no one type of graph is equally good for all purposes. The histogram is the easiest of all to understand and is usually the best if only one distribution is being represented. If however two or more distributions are to be compared, polygons are usually better, since so many lines coincide when histograms are superimposed that the picture is likely to be confusing. A frequency polygon is less precise than a histogram in that it does not represent accurately i.e., in terms of area, the frequency upon each interval. Both the frequency polygon and the histogram tell the same story and both enable us to show in graphic form how the scores in the group are distributed, whether they are piled up at the low or high end of the scale or are evenly distributed over the scale. But when it comes to comparing different groups of unequal sizes, both the above graphs present difficulties. In such a case, the percentile curve is extremely useful. It permits comparisons regardless of the sizes of the groups. A discussion of this and other types of graphs will be found in books dealing mainly with statistics.

It should be noted, however, that there are better methods of comparing groups than graphic methods, methods which are precise and which also take into account chance differences between groups We shall refer to these methods in the succeeding chapters.

Smoothing Frequency Polygons

When a frequency polygon is plotted with a limited number of cases, the outline of the polygon is usually irregular. Many of these irregularities may be considered as only accidental or of little or no significance, since they may be peculiar to the particular grouping employed in the frequency table, or may be characteristic only of one sample of individuals considered and not generally characteristic of other similar groups. We frequently wish to forecast, from the supposedly representative sample that we have, how a larger population would distribute itself. We would like to know how the polygon would look like if the grouping errors and sampling errors were removed from the polygon.

In order to obtain a more highly generalized picture, therefore, the practice of 'smoothing' the original figure is sometimes followed. In smoothing a polygon, we try to ascertain the shape which it would take if it represented conditions freed from minor accidental fluctuations. There are various methods, but one of the better methods of smoothing a distribution is that of "running or moving averages". In this method, a series of running or moving averages are taken from which new or adjusted frequencies are determined (Table 14.5).

Running Averages : The process of taking moving averages by 3's is illustrated below. The same data as given in Table 14.3

Table 14.5 : Original and smoothed frequencies for the distribution given in Table 14.3.

Scores X	Observed frequency	Smoothed frequency
140-149	0	$\frac{0+0+2}{3} = 0.67$
130-139	2	$\frac{0+2+7}{3} = 3.00$
120-129	7	$\frac{2+7+11}{3} = 6.67$
110-119	11	$\frac{7+11+14}{3} = 10.67$
100-109	14	$\frac{11+14+20}{3} = 15.00$
90- 99	20	$\frac{14+20+16}{3} = 16.67$
80- 89	16	$\frac{20+16+7}{3} = 14.33$
70- 79	7	$\frac{16+7+3}{2} = 8.67$
60- 69	3	$\frac{7+3+0}{3} = 3.33$
50- 59	0	$\frac{3+0+0}{3} = 1.00$
	80	80.00

are used. It will be noted that two zero frequency classes are appended at the top and bottom of the observed distribution.

The smoothed frequency for the interval 90-99 is obtained by adding the frequencies in the classes just above and just below that class and dividing by 3. Thus (14 + 20 + 16) / 3 = 16·67 is the smoothed f; for interval 120–129, it is (11 + 7 + 2)/3 = 6·67. The smoothed f's for the other intervals may be found in the above table. To find the smoothed f's for the two intervals at the extremes of the distribution, namely, 60–69 and 130–139, we add two zero frequency classes. The smoothed 'f' for 60–69 is then (0 + 3 + 7)/3 = 3·33 and that for 130–139 is (7 + 2 + 0)/3 = 3·00. The smoothed 'f' for the interval 50–59 for which the 'f' in the original distribution is O, is (3 + 0 + 0)/3 = 1·00 and that for the interval 140–149, it is $\frac{0+0+2}{3}$ = 0·67. Note that if we omit these last two intervals, the N for the smoothed distribution will be less than 80, as the original smoothed distribution has f's outside the range of the original distribution (Table 14.5).

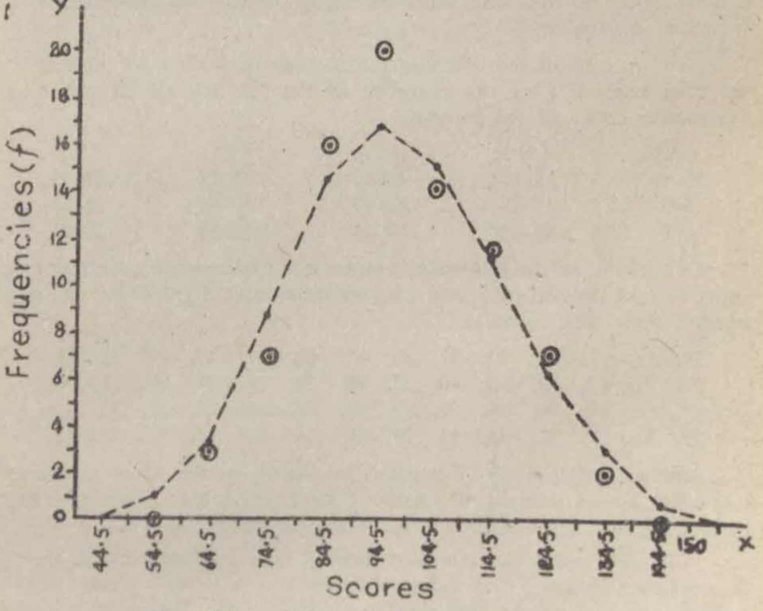


Fig. 14.4 : A Smoothed Frequency Curve.

Plotting a Smoothed Distribution : The final step is to plot the smoothed curve which we have in Fig. 14.4. First the obtained frequencies are plotted as circles in their proper places. The smoothed frequencies are then plotted as points and a smooth curve is drawn. In doing this, the curve need not necessarily pass through all these points. When an investigator presents only the smoothed frequency polygon or a smoothed curve and does not give his original data, it is impossible for a reader to tell with what he started. Moreover, smoothing gives a picture of what an investigator might have got, if his data had been more numerous, or less subject to error than they were. Hence it is well to indicate observed frequencies along with the smoothed ones. This is a fundamental point to keep in mind in all curve plotting.

FOR FURTHER STUDY

(1) How can statistics be useful to a teacher, in understanding the performance of her students in a certain subject ? Illustrate.

(2) Illustrate the various ways of organizing the scores in the form of a frequency distribution. What are the advantages and limitations of a frequency distribution ?

(3) In each of the following distributions, indicate (a) the size of the class interval, (b) the mid-point of the intervals shown and (c) the precise limits of the intervals.

(A)	(B)	(C)	(D)	(E)
10–14	18–20	48–55	70–79	24–25
5–9	15–17	40–47	60–69	22–23
0–4	12–14	32–39	50–59	20–21

(4) Tabulate the following 56 scores into two frequency distributions, using (i) an interval of 3, and (ii) an interval of 5 units. Let the first interval begin with score 15.

59, 43, 37, 34, 30, 22, 29, 56, 42, 37, 33, 29, 22, 21,
32, 36, 42, 52, 50, 40, 32, 36, 32, 28, 21, 20, 27, 36,
39, 50, 47, 38, 36, 31, 27, 27, 17, 46, 38, 35, 31, 17,
25, 31, 35, 38, 44, 43, 29, 37, 34, 30, 24, 38, 35, 41,

Supply the following information by inspecting the above frequency tale which has an interval of 5 units. [Minimum score for getting a first class is 60, for a second class it is 50 and for passing it is 35.]

[i] How many students have failed ? [ii] What percent of the students have obtained first and second class ? [iii] What percent of the students obtained scores between 20 and 29 ? [iv] What can you say

about the general performance of this class ? [v] If this class is known to be an above average class, what inference can be drawn about this particular test ?

(5) [a] Plot a frequency polygon of the 56 scores given in (4) above, using an interval of 3 units. Superimpose a histogram upon the frequency polygon, using the same axes.

[b] Plot a frequency polygon of the 56 scores given in (4) above, using an interval of 5 units. Smooth the distribution and draw a smoothed curve. Plot on the graph both the observed and smoothed frequencies. In what respects has smoothing changed the picture of these data ?

(6) Following are the scores of two groups on a test in arithmetic Compare the two distributions by plotting frequency polygons. Which of these two groups has a greater variability ?

(For a further comparison of these groups by statistical methods, refer to question 11 at the end of chapter 16).

Scores	A	B
80–89	0	5
70–79	1	6
60–69	5	5
50–59	10	7
40–49	20	7
30–39	15	9
20–29	7	10
10–19	2	6
0– 9	0	5
	N=60	N=60

(7) *Answer the following questions :*

[i] What is the meaning of a score such as 65, 66 of 67 in a continuous series ?

[ii] Explain the various steps in preparing a frequency distribution from raw scores.

[iii] How does the size of class intervals affect a frequency table ?

[iv] What are the advantages of graphical representation of test scores ?

[v] In what respects does a frequency polygon differ from a histogram ?

[vi] Why is it customary to smooth a frequency polygon ? Why is it necessary to show the observed frequencies along with the smoothed ones ?

(8) *Explain the following statements* :

[a] In a frequency distribution, the individual scores lose their identity.

[b] A frequency polygon is better than a histogram when two or more distributions are to be compared.

[c] You cannot easily compare two groups of unequal sizes with the help of a histogram or a frequency polygon.

[d] Statistics in the hands of a beginner is an uncertain tool and in the hands of a propagandist, a dangerous one.

[e] In making a frequency distribution table, something is lost and something is gained.

(9) Prepare a frequency table for the following test scores using a class interval of 3. Begin the first interval from 4.

23, 20, 18, 13, 12, 7, 4, 8, 11, 15, 16, 15, 14, 10,
10, 13, 11, 14, 17, 14, 14, 17, 14, 9.

(10) Miss Medha's class had the following scores on a test in science
37, 38, 36, 31, 28, 33, 24, 19, 25, 34, 16, 43, 22, 20, 26,
44, 27, 19, 25, 34, 33, 24, 22, 20, 44, 27, 31, 28, 38, 17,
31, 26, 34, 17, 19, 20, 22, 24, 26, 29.

Set up a frequency table using an interval of 5 units. Let the first interval begin from 10. [Minimum score for getting a first class is 50, for a second class 40 and for passing 20]. What inference would you draw about the class from the performance in this test ?

* * *

15

MEASURES OF CENTRAL TENDENCY

The grouping of test scores into frequency tables is one step in the process of condensing them so that they can be analysed and interpreted. However, a further step must be taken before it is possible to describe the data. Some single term or value that is representative of the entire table must be found. Since these values which may be taken to represent an entire distribution of scores are usually found near the centre of the data when arranged in order of size, they are commonly called *measures of central tendency*.

When the scores are classified into frequency distributions or are graphically depicted in histograms or frequency curves, points of similarity and difference between the various curves may be noted roughly by inspection. But comparisons by inspection tend to be inexact and inconclusive, and it is difficult to obtain agreement regarding their meaning. As a rule quantitative methods of comparing frequency distributions are more satisfactory than graphical methods. There are four different ways in which frequency distributions may differ. They are : (i) average value or central tendency, (ii) scatter or variability, (iii) skewness and (iv) kurtosis. Two or more comparable groups generally differ to some extent in all these four respects. Hence to compare two or more groups we need measures of these four characteristics. These measures in addition to providing the basis for exact comparisons of two series, are indispensable in the analysis and interpretation of a single series. We shall consider only the measures of central tendency in this chapter. There are five measures of central tendency : the arithmetic mean, the median, the mode, the geometric mean and the harmonic mean. The first three are by far the most widely used measures, hence our discussion will be limited to these three. The value of a measure of central tendency is two fold. First, it is an average, which represents all of the scores made by a group, and as such gives a concise description of the performance of the group as a whole, and second it enables us to compare two or more groups in terms of typical performance. To avoid confusion, think of an average as a point on the scale and not as a score of any particular individual. Sometimes though, it may coincide with the actual score.

1. The Mean

(a) *Calculation of the Mean when data are ungrouped* : The arithmetic mean or simply the mean is the best known measure of central tendency. Indeed the word 'average' is thought by many persons to designate the arithmetic mean, although the arithmetic mean is only one of several averages. Most readers already know that to find the arithmetic mean, we sum the measurements and then divide the total by the number of cases. The formula for the mean (M) of such a series is :

$$M = \frac{\Sigma X}{N} \quad \text{(arithmetic mean) .... (1)}$$

in which N is the number of measures, X stands for a score and the symbol Σ means 'sum of'. The symbol Σ will occur repeatedly in subsequent work and should be clearly understood. It is not a separate quantity. It should be read, 'the sum of' what follows. ΣX means the sum of all X numbers. Table 14.1 gives a list of scores on an intelligence test. The mean score for the group can be obtained by adding all the scores and dividing the sum by the number of students.

$$M = \frac{\Sigma X}{N} = \frac{7896}{80} = 98.7$$

This method will not be useful when the data are in the form of a frequency table, or when the series is too lengthy for comfortable addition.

(b) *Calculation of the Mean from grouped data (Long Method)* : When scores are put into a frequency distribution, the scores classified within any interval lose their identity and are represented by the mid-point of that interval. This necessitates a slightly different procedure from the one given above and used with ungrouped scores. Illustration in Table 15.1 will make the difference clear. The same scores given in Table 14.1 and 14.3 are considered.

Table 15.1 : Calculation of the Mean from grouped data.

1 Scores	2 Mid-point of the interval	3 Fre-quency	4 Product	Procedure
X	X_m	f	$f. X_m$	
140-149	144.5	0	0.00	(i) Write down the mid-point of each class-interval in column 2, (X_m). (ii) Mid-point is then multiplied or weighted by the frequency lying opposite it in the 'f' column and the product is written in the fourth column. (iii) Sum of the $f.X_m$ column divided by N gives the value of the mean. $M = \frac{\sum f. X_m}{N}$ ... ... (2)
130-139	134.5	2	269.00	
120-129	124.5	7	871.50	
110-119	114.5	11	1259.50	
100-109	104.5	14	1463.00	
90- 99	94.5	20	1890.00	
80- 89	84.5	16	1352.00	
70- 79	74.5	7	521.50	
60- 69	64.5	3	193.50	
50- 59	54.5	0	0.00	

N = 80 $\sum f. X_m = 7820.00$

$\therefore$ Mean $= \frac{\sum f. X_m}{N} = \frac{7820.00}{80} = 97.75$

The mean as calculated from the same but ungrouped scores comes out to be 98·7. The slight discrepancy between the mean of grouped data and that of ungrouped data is due to errors of grouping. We assume that the mid-point of an interval correctly represents all the scores within that interval. This will not be true in many cases and to that extent errors of grouping will be present. We may also expect minor discrepancies in means when the same data are grouped differently, that is, with different sizes of class-interval or with different starting points for intervals of the same size.

Remember that if the class interval say of 120–129 is considered as extending from 120 to 130 as shown in Table 14.4 (A), the mid-point of each class interval will be shifted by 0·5 of a score point and this would result in an arithmetic mean exactly 0·5 larger than the value we have obtained in Table 15·1.

(c) *Computation of the Mean when the distribution has a unit interval*: The example given in Table 15.2 differs from the previous one in that the interval is of one unit only. Hence there is no necessity of having the column 'mid-point of the class interval.' There are 8 students having a score of 16, hence their contribution is 16 times 8, which is shown in the third column of Table 15.2. Otherwise the procedure is the same as described in the previous example.

Table 15.2: Computation of the Mean, when the distribution has a unit interval.

Score (X)	Frequency (f)	$f.X$
19	1	19
18	1	18
17	3	51
16	8	128
15	17	255
14	13	182
13	9	117
12	6	72
11	0	0
10	2	20
	$N = 60$	$\Sigma f.X = 862$

$$\text{Mean} = \frac{\Sigma f.X}{N} = \frac{862}{60} = 14 \cdot 37$$

(d) *The 'Short Method' of Computing the Mean*

The long method described above is a straightforward one: it gives accurate results but often requires the handling of large numbers and entails tedious calculations. Because of this, the *Assumed Mean Method* or simply *the Short Method* has been devised for computing the mean. This method involves the principle of a deviation score. If we substract the mean from any score, the result is the deviation of that score from the mean. Symbolically :

$$x = X - M$$

where X is any individual score
M is the mean of the group and
x is the deviation of the score from the mean.

Consider the scores given below; the mean of the series is 6.

											Total	M
Raw scores :	3,	8,	9,	7,	4,	7,	3,	4,	9,	6	60	6
$x = X - M$:	−3	+2	+3	+1	−2	+1	−3	−2	+3	0	0	

Take the deviation of each score from the mean, taking into consideration the signs. The sum of all these deviations is zero. This is an important property of the mean, namely, the algebraic sum of the deviations of all the scores from the mean is zero. Consider the same ten scores (Table 15.3).

Table 15.3 : Short method of determining the Mean.

Raw Scores X	Assumed Mean AM	Deviation x'=X−AM	Procedure
3	5	− 2	(i) Let us assume the mean to be 5 (ii) Find the deviation of each score from the assumed mean : x'=X−AM (iii) Add the deviations and divide it by N $\frac{\sum x'}{N} = \frac{+10}{10} = +1$; this is the correction c. (iv) Add the correction to the assumed mean True mean M=AM+c=5+1 .... (3) Check this value with the one obtained previously.
8	5	+ 3	
9	5	+ 4	
7	5	+ 2	
4	5	− 1	
7	5	+ 2	
3	5	− 2	
4	5	− 1	
9	5	+ 4	
6	5	+ 1	
		$\sum x' = +10$	

The above characteristic of the mean provides another method for computing the mean. We can guess what the mean of a distribution is likely to be and then correct our guess for the true mean. Let is use this method for finding out the mean of scores in Table 15.4.

Table 15.4 : Computation of the Mean by the short method

Scores X	Midpoint of the interval X_m	Frequency f	Deviation $x' = X_m - AM$	Product $f.x'$
140–149	144.5	0	0	0
130–139	134.5	2	4	8
120–129	124.5	7	3	21
110–119	114.5	11	2	22
100-109	104.5	14	1	14
				+ 65
90– 99	94.5	20	0	0
80– 89	84.5	16	–1	–16
70– 79	74.5	7	–2	–14
60– 69	64.5	3	–3	–9
50– 59	54.5	0	–4	0
				–39
AM = 94·5		N = 80		Σfx' = + 26

Correction in interval units $c = \frac{\sum f.x'}{N} = \frac{+26}{80} = +0.325$ .... (4)

Correction in score units $= i \times c = 10 \times (\cdot 325) = 3\cdot 25$

True mean $M = AM + ci = 94\cdot 5 + 3\cdot 25 = 97\cdot 75$

Steps in Computing the Mean

(i) Assume a value for the mean. Choose the mid-point of an interval near the middle of the distribution. Here we select the interval 90-99 and its mid-point 94·5 is the assumed mean AM.

(ii) Call this assumed mean (AM) zero, and denote it in the x' column. In this column lay off the deviations from the assumed mean by intervals. Thus 104·5, the mid-point of 100–109 deviates from 94·5 the AM, by one interval, and hence '1' is placed in the x' column opposite 104·5. In like manner 114·5 deviates two intervals from the AM and a '2' goes in the x' column opposite 114·5. Similarly counting downwards, show the deviations of the scores 84·5, 74·5, by –1, –2 in the x' column.

[x' is used to denote the deviation of a score X from the assumed mean (AM), while x is the deviation of a score X from the actual mean.]

(iii) Multiply the x' (deviation) values by their frequencies and enter the products in the fx' column. All fx' on intervals above (greater than) the AM are positive and all fx' on intervals below (smaller than) AM are negative. Determine the algebraic sum of the deviations of the scores in the fx' column. This sum comes out to be + 26 in Table 15.4.

(iv) Divide the algebraic sum of the deviations by N. This gives the correction c (+ 0·325) in units of class interval. If we multiply c by i (the size of the class interval), the result is the correction (3·25) in score units. Add this correction to 94·5, the AM; the result is the actual mean, 97·75. See that the value is the same as obtained by the long method in Table 15.1.

The correction may sometimes come out to be zero. In that case, it would mean that the assumed mean is the same as the true mean. This method is sometimes called the *coding method,* since the x' units really are coded values of the mid-points of the class intervals. Remember that we can select the new origin (AM) anywhere in the distribution, though in order to obtain the greatest benefit from the coding method, it is well to choose the origin near the centre of the distribution.

(e) *The Mean from Combined Samples or Groups*

On a vocabulary test the mean for a group of 15 children is 40 and the mean on the same test for another group of 35 children is 54. What will be the mean of the two groups combined? The formula for the weighted mean of n groups is:

$$M = \frac{N_1 m_1 + N_2 m_2 + N_3 m_3}{N_1 + N_2 + N_3} \quad \ldots\ldots\ldots\ldots (5).$$

Where M = mean of the combined groups

N_1, N_2, N_3 etc., are the number of cases in each group

m_1, m_2, m_3 etc., are the means of the respective groups.

When only two groups are combined, the formula becomes

$$M = \frac{N_1 m_1 + N_2 m_2}{N_1 + N_2}$$

Substitute the values as given in the above example

$$M = \frac{(15 \times 40) + (35 \times 54)}{15+35} = \frac{2490}{50} = 49.8$$

2. The Median

A much more useful way of representing the typical or average score is to find the value on the score scale that separates the top half of the group from the bottom half. This is called the *median*. It is defined as that point on the scale of measurement above which are exactly half the cases and below which are the other half. Note that it is defined as a point and not as a score or any particular measurement. In an ungrouped series, the term 'mid-score' or 'mid-measure' is used, while the term 'median' is used only in the case of a grouped frequency distribution. The mid-measure is designated as an actual score while the median is a point on the scale; it does not show any individual's actual score.

(a) *Finding the Mid-score from ungrouped data*

When ungrouped scores or other measures are arranged in order of size, score in the middle of the series is the mid-score or mid-measure. Two situations arise in the computation of the mid-measure from ungrouped data: (i) when N is odd and (ii) when N is even. Consider first, the case where N is odd:

9 14, 8, 12, 9, 11, 10.

Arrange these seven scores in order of size:

8, 9, 9, 10, 11, 12, 14,

the mid-score is 10, because it is exactly the middle score in the series. There are three scores below 10 and three scores above 10. Now consider the series when N is even:

8, 9, 18, 21, 25, 33

Strictly speaking when N is even, there is no mid-score. In that case it is customary to average the middle pair of scores. The two middle scores are 18 and 21. The mid-score is half way between them. Hence the mid-score is $\frac{18 + 21}{2} = 19{\cdot}5$

A formula for finding the mid-score from ungrouped data is: Mid-score $= \left(\frac{N+1}{2}\right)^{th}$ measure in order of size. In the first illustration above, the mid-score is the 4th score counting in from either end of the series, which is 10. In the second illustration the mid-score is on the $\left(\frac{6+1}{2}\right)^{th}$ or 3 · 5th score in order of size, that is 19·5

$$\text{Mid-score} = \left(\frac{N+1}{2}\right)^{th} \text{ measure in order of size} \ldots\ldots \quad (6)$$

(b) *Computing the Median of grouped data*

When scores in a continuous series are grouped into a frequency distribution, the median by definition is the 50% point in the distribution. To locate the median, therefore, we take 50% (i.e. N/2) of our scores and count into the distribution until the 50% point is reached. In the illustration in Table 15.5, there are 80 cases and hence the median is that point on the measuring scale above which there are 40 cases and below which there are 40. Starting with the lowest score, we count up until we have the necessary 40 cases. Table 15.5 shows the cumulative frequencies as well as the frequency in each interval. Cumulative frequency shows the total number having a score equal to or less than the highest score in that interval. Thus there are 10 cases scoring at or below 79·5 and 26 cases scoring at or below 89·5. As indicated, we wish to identify the point below which 50 percent of the cases fall. Since 50 percent of 80 is 40, (N = 80), we must identify the point below which 40 pupils fall. We note that 26 pupils have scores below 89.5. We need to include 14 more cases to make up the required 40 cases. Hence the median lies some where within the interval 90–99, whose exact limits are 89·5–99·5. We assume for the sake of computation that the 20 cases within this interval are

Table 15.5. Computing the Median of Grouped Data.

(1) Scores X	(2) Frequency f	(3) Cumulative Frequency	Procedure
130-139	2	80	N = 80 N/2 = 40 Median falls in the interval 90–99 Lower limit of this interval $l = 89.5$ Number of cases below the median interval, F = 26 Size of the class interval $i = 10$ Frequency within the median interval, $f = 20$
120-129	7	78	
110-119	11	71	
100-109	14	60	
90- 99	20	46	
80- 89	16	26	
70- 79	7	10	
60- 69	3	3	

$$Mdn = l + \left(\frac{N/2 - F}{f}\right) \times i = 89.5 + \left(\frac{40 - 26}{20}\right) \times 10$$

$$= 89.5 + 7 = 96.5 \ldots\ldots\ldots\ldots (7)$$

evenly spread over the distance from 89.5 to 99.5 (see fig. 15.1). In order to get the 14 extra scores needed to make exactly 40, we must go 14/20 of the way, for 14 is the number we require and 20 is the total number in the interval. The size of the interval is 10 units and so on the scale of measurement we go 14/20 × 10 or 7 units and add this to 89.5, the lower limit of the median interval. This puts the median at 89.5 + 7.0 = 96.5.

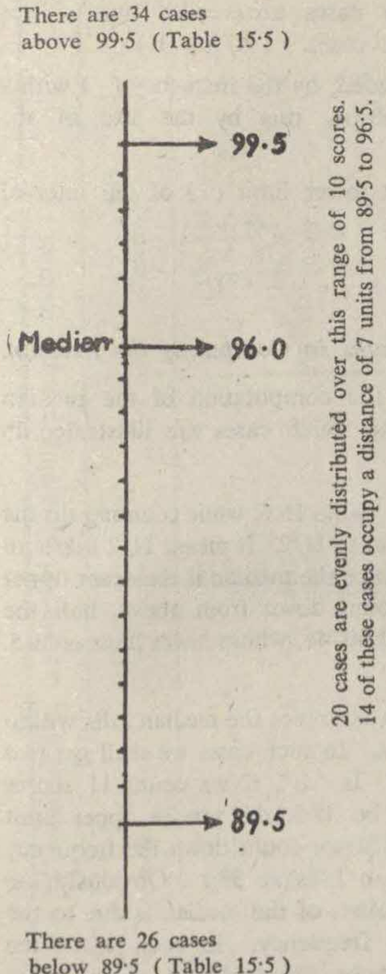

There are 34 cases above 99·5 (Table 15·5)
99·5
Median
96·0
89·5
20 cases are evenly distributed over this range of 10 scores.
14 of these cases occupy a distance of 7 units from 89·5 to 96·5.
There are 26 cases below 89·5 (Table 15·5)

Fig. 15.1. Enlargement of the Median Interval.

We can check this by counting down from the top of the distribution until we include N/2 of the cases, 40 in this problem. Starting at the top, we find that $2 + 7 + 11 + 14 = 34$ we need six more cases out of the next group of 20. We must go 6/20 of the way below the upper limit of the interval i.e., below 99·5. This means 6/20 × 10 or exactly 3·0 units. Thus the median is $99{\cdot}5 - 3{\cdot}0 = 96{\cdot}5$, which checks with the one obtained by counting from below. It is always a good practice to check the value of the median in this manner. If the two estimates do not agree exactly, something is wrong. Apply the same procedure for finding out the median when the distribution has unit interval, as given in Table 15·2.

Steps involved in interpolating the Median

(1) Find N/2 or half the number of cases in the distribution.

(2) Count up from the lower end until the interval containing the median is located. (F)

(3) Determine how many cases are needed out of this interval to make N/2 cases. (N/2 - F)

(4) Divide this number needed, by the frequency (f) within the interval and multiply this by the size of the interval. (i).

(5) Add this to the exact lower limit (l) of the interval containing the median.

$$Mdn = l + \left(\frac{N/2 - F}{f}\right) \times i \;\ldots\ldots\ldots\ldots\; (7)$$

(c) *Some Special Situations in Computing the Median.*

There are cases in which the computation of the median presents certain unusual features. Such cases are illustrated in Table 15.6.

(a) In the first example in Table 15.6, while counting up the cases from below, 'F' is just equal to N/2. It means N/2 takes all the cases in the interval 35–39, hence the median is the exact upper limit of that interval. If we count down from above, half the cases take us through the interval 40–44, whose lower limit is 39·5. Hence the median is 39·5.

(b) and (c): In these two examples the median falls within an interval having zero frequency. In such cases we shall get two different values for the median. In 'b', if we count 11 scores from below, the median would be 35·5, the precise upper limit of the class interval 33–35, while if we count down the frequency distribution 11 scores, the median falls at 38·5. Obviously the discrepancy between these two values of the median is due to the interval 36–38 which has zero frequency. In such cases the median is usually considered to be the midpoint of the class interval having zero frequency. In illustration (b) the median is considered to be 37·0 rather than 35·5 or 38·5, while in (c) the median is 11·5.

(d) Illustration 'd' differs from most cases in that the frequency of the lowest class is greater than N/2. In such cases l of the formula is the lower limit of the bottom class and F, the number of cases below the median interval is zero. The median comes out to be 25·5.

Table 15.6. Computation of the Median presenting unusual features.

(a)		(b)		(c)		(d)	
Scores	f	Scores	f	Scores	f	Scores	f
55-59	5	45–47	1	20–21	2	130–149	1
50-54	3	42–44	4	18–19	1	110–129	3
45-49	8	39–41	6	16–17	0	90–109	0
40-44	18	36–38	0	14–15	3	70–89	2
35-39	15	33–35	7	12–13	0	50–69	4
30-34	10	30–32	3	10–11	0	30–49	5
25–29	7	27–29	1	8–9	2	10–29	25
20–24	2			6–7	2		
				4–5	1		
				2–3	1		
	$N = 68$		$N = 22$		$N = 12$		$N = 40$
	$N/2 = 34$		$N/2 = 11$		$N/2 = 6$		$N/2 = 20$
$Mdn = l + \left(\frac{N/2-F}{f}\right) \times i$ $= 34{\cdot}5 + \left(\frac{34-19}{15}\right) \times 5$ $= 39{\cdot}5$			$Mdn = 37{\cdot}0$		$Mdn = 11{\cdot}5$	$Mdn = l + \left(\frac{N/2 - F}{f}\right) i$ $= 9{\cdot}5 + \left(\frac{20 - 0}{25}\right) \times 20$ $= 9{\cdot}5 + 16$ $= 25{\cdot}5$	

(*d*) *The Median in a Histogram*

As was seen in chapter 14, the area of the histogram is proportional to the number of scores in a distribution. Since the median by definition is the point on the scale of scores below (or above) which one half of the scores lie, a vertical line drawn through the median will bisect the histogram, and conversely, the vertical line which bisects the histogram will pass through the median.

3. The Mode

The mode may be defined as the item which occurs most frequently in a statistical series. When a particular type of dress is worn more frequently than other types in a particular season, that particular type of dress is referred to as the mode for that season. In a simple ungrouped series of measures, mode is that single measure or score which occurs frequently. For example, in the series 6, 8, 9, 10, 11, 11, 11, 13, 14, 14, 15 the most often recurring measure namely 11, is the mode. When data are grouped into a frequency distribution the mode is usually taken to be the midpoint of the interval having the largest frequency. In Table 15.5, the interval 90–99 contains the largest frequency and hence 94·5, its midpoint, is the mode.

The mode obtained by mere inspection, as suggested above is called the *crude, apparent,* or *inspectional* mode. It is almost instantaneously obtainable and it frequently serves as a good typical value of the thing being measured.

A distribution may have more than one mode. If there are two intervals with the same maximum frequency, and if they are separated by more than one intervening interval of lower frequency, we should be justified in saying that the distribution is *bimodal.* This is seen in group B in Table 15.7. An arithmetic test is given to four groups and the score distributions for the four groups are given in Table 15.7. Group A has a marked mode at 52, the midpoint of the interval 50–54. In group B, there are two outstanding frequencies and it has two modal values. We are justified in saying that the distribution is bimodal (Fig. 18.8). Such a situation may occur when a class consists of two

Table 15.7 : Determining the Crude Mode for four groups.

Arithmetic Scores	Group A	Group B	Group C	Group D
80–84	0	2	2	2
75–79	1	3	4	10
70–74	3	15	6	6
65–69	4	2	12	10
60–64	6	6	12	4
55–59	7	0	4	5
50–54	16	6	3	3
45–49	5	15	0	10
40–44	4	4	3	9
35–39	4	3	2	1
	N = 50	N = 56	N = 48	N = 60

distinct types of pupils such as very bright and very dull students. In group C the two maximum frequencies occur in adjacent class intervals. There are not two distinct groups in this class; the distribution is therefore probably really *unimodal* and the crude mode will be the dividing point between these intervals which is 64·5. That this distribution is unimodal, will be seen, if we increase the size of the class interval from 5 to 10 units. The distribution in D above, appears to have several modes. None of the modes are prominent enough, however, to support any conclusions about single modal values. Such distributions possess no marked mode and when dealing with them, we cannot use modal values either in description or in making inferences about the central tendency of the group. It appears that group D consists of pupils of varying abilities and they appear to cluster about 3 or 4 levels. Such a distribution is called *multimodal.* The presence of two or more modal values in a distribution casts doubt upon the homogeneity of the group. It suggests that important qualitative differences may underlie the data and that more exact and useful comparisons may be possible, if the differences are taken into account. Such a situation also invites questions regarding the conditions giving rise to such data.

The True Mode : When calculating the mode from a frequency distribution, we distinguish between the *true mode* and the *crude mode.* The true mode is the point (or peak) of greatest concentration in the distribution; that is the point at which more measures fall than at any other point. When the scale is divided into finely graduated units, when the frequency polygon has been smoothed and when N is large, the crude mode approximates the true mode. The true mode cannot actually be determined but an approximation to it, known as the *estimated, computed,* or *refined* mode is often found. Various formulae for this estimation have been suggested; we shall consider here only one of them. This formula is applicable when the frequency distribution is symmetrical or at least not badly skewed. The formula is :

$$\text{Mode} = 3\ Mdn - 2\ \text{Mean} \qquad \ldots \quad (8)$$

[The distribution in example 8 (h) at the end of this chapter is badly skewed, hence this formula is not applicable there.]

Applying this formula to distribution in Table 15.1 and 15.5, we estimate the mode of the distribution :

$$\text{Median} = 96{\cdot}5, \qquad \text{Mean} = 97{\cdot}75$$

$$\therefore \text{Mode} = 3\ (96{\cdot}5) - 2\ (97{\cdot}75) = 94{\cdot}00$$

By inspection we get the crude mode as 94·5. Thus the two are approximately equal.

What are the characteristics of the Mean, the Median and the Mode ? : Because of the difference in the method of determining these measures, mean is often called the *calculated* average, median is known as the *counting* average while the mode is called the *inspection* average.

The Mean

(1) The concept of the mean is very familiar. It can be accurately determined with the help of different methods. (2) Mean is better suited to further arithmetical computations such as the standard deviation and correlation. (3) Mean is algebraic in nature, that is the means of separate groups can be combined to give the mean of the combined group. (4) The mean is very sensitive to the size of the extreme scores. Consider the following example :

							Mean	Median
5	6	9	11	14	18	21	12	11
5	6	9	11	14	44	65	22	11

It will be seen that the mean is affected by the size of a few extreme cases while the median remains the same. (5) Mean has an important property namely, the sum of the deviations of all scores from the mean is zero, (see page 291). The median does not qualify in this respect; consider the deviation of the scores from the median in the series given above. The algebraic sum of the deviations is 7·0. If the fulcrum of a bar is placed at the median, it will not be perfectly balanced. This comes from the fact that in computing the median we ignore the distance of each score from the central value.

Limitations of the Mean : (i) Though the mean is the most reliable measure, it loses this advantage as the distributions become more and more skewed. For very skewed ones, it may be even less reliable than the median or mode. Any extreme score at one end of the series, will affect the average and thus tend to make it unrepresentative. (ii) It cannot be accurately determined in the case of an open-end table; a table that does not give the exact scores of the extreme 4 or 5 cases is called *an open end table*. (iii) Unless the data are very simple, the average cannot be obtained by inspection. (iv) Mean cannot be used with scores expressing rank or order of students such as A, B, C, D . . . . . . R, S. T. etc. It can be used only with distributions that give absolute scores.

The Median

(1) The median is easily understood and has several advantages as an average. When a series contains either a few

extremely high or extremely low scores, the median is perhaps the most representative average available. Suppose a factory has 100 employees; five of them have their monthly salaries of Rs. 1000 each while the rest have their salaries between Rs. 50 and Rs. 200. The mean salary for the factory is likely to be misleadingly high. However, the median will not be sensitive to the few high salaries. In fact, the median can be ascertained without even knowing the actual salaries of the five top employees. (2) When the central tendency of an open end distribution i.e., a distribution having a bottom or top interval of unspecified length is desired, the median is the most reliable measure that can be computed. (3) Median can be used not only for distributions giving absolute scores such as 60, 65, 71, 84 etc., but also for distributions giving ranks such as A, B, C, D, E etc. When the units of measurement are unequal, median is preferable; it does not assume equality of units, whereas the mean assumes it.

Limitations of the Median : (1) The data have to be arranged in an ascending or descending order, before the median can be located. (ii) It cannot be used for computing other statistical measures such as the standard deviation and the correlation coefficient. (iii) The median like the mode is a non-algebraic measure and medians of separate distributions cannot be combined to give the median of the combined distribution. (iv) It cannot be used when due weightage is to be given to all the scores.

The Mode

(1) The mode which indicates the item of greatest frequency in a series, is an average widely used in everyday life. When newspapers use the term *on the average*, they are usually referring to an outstanding tendency and not an arithmetic average. (2) Mode is an average which is easy to understand and easy to determine by mere inspection. (3) It is not affected by extreme scores and we need not know the magnitude of the extreme scores in a series to determine the modal score.

Limitations of the Mode : (1) The mode is affected by changes in the grouping scheme. Consider the distribution given in Table 15.7. If we change the size of the class interval from 5 units to 10, the modal values will fluctuate. (2) It is not a very trustworthy measure especially with small groups, for example, in the series 17, 27, 27, 27, 32, 43, 48, 48, 49, 50, the modal value is 27, because more pupils obtained that score. In this case, the changing of two scores might shift the mode decidedly. If one of the pupils who made 27 had scored 28, and if the one who had made 49, had scored 48, the mode would suddenly shift from 27 to 48 since more pupils would then have made that score than any other. (3) In ungrouped data, some-times it is difficult to determine the mode, simply because no score is repeated e.g., 7, 10, 12, 15, 21, 23, 42, 46. Does it mean that there is no modal value ? (4) The mode like the median, suffers from the additional disadvantage of being incapable of algebraic treatment. The mode of one distribution cannot be combined with the mode of another, to determine the mode of the combined distributions.

Central Values in Skewed Distributions

In a normal curve, the three measures coincide. Even in symmetrical distributions (which may not be necessarily normal) the mean and the mode coincide with the middle point or the median. But as the frequency distribution tends to be asymmetrical, the mean and the mode separate from the median. As the median divides the distribution into two equal parts, it will always be in the middle, and the other two measures, namely the mean and the mode will be on opposite sides of the median in a skewed

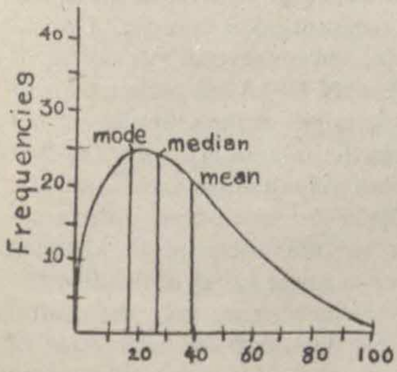


Fig. 15.2 : Positive skewness

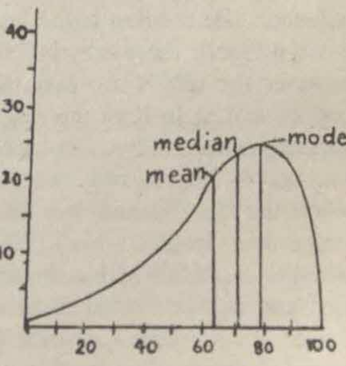


Fig. 15.3 : Negative skewness.

distribution. The mean is always pulled towards the skewed (pointed) end of the curve (Fig. 15.2, 15.3). The arithmetic mean, as the centre of gravity of the distribution is weighted toward the extreme values. In skewed distributions, it is usually well to state both the mean and the median, since each tells its own story and from the difference between the two, we can infer about the direction of skewness. For further discussion of skewed distributions, refer to chapter 18.

When to use the various measures of central tendency

As a rule the mean is regarded as the best measure of central tendency and the crude mode as the poorest. But there are occasions when one is better suited than the other.

Use the Mean : (i) When the greatest reliability is wanted. It usually varies less from sample to sample drawn from the same population. (ii) When other statistics (e.g., standard deviation, coefficient of correlation) are to be computed later; many statistics are based upon the mean. (iii) When the distribution is symmetrical about the centre. (iv) When we wish to know the centre of gravity of a sample, that is, when the contribution of each score is to be considered.

Use the Median : (i) When the exact mid-point of the distribution is wanted. (ii) We have seen that the mean is greatly influenced by extreme scores, and whenever we wish to avoid this influence, the median is the best measure. For example, if a test is too difficult, there may be several zero or several poor scores; if however the test is too easy there may be several perfect or near perfect scores; in both the cases the pupils at the extremes are not correctly measured. In such cases the median is usually the best average to use. (iii) The median is a more reliable measure, when the distribution has an upper or lower class interval of unspecified length. (iv) When we are interested in knowing whether cases fall within the upper or lower halves of the distribution and not interested in knowing the distance from the central point. (v) When a distribution is described and interpreted in terms of percentiles; median is a member of the percentile system and hence is the appropriate measure in such a situation.

Use the Mode : (i) When a quick and approximate measure of central tendency is wanted. (ii) When information regarding the most typical case, such as the style of dress or shoes is all that is wanted.

FOR FURTHER STUDY

(1) Describe three specific situations in which either the mean, the median or the mode may be most useful. Under what conditions would it be desirable to use more than one measure in summarizing the data ?

(2) [a] Draw a frequency polygon each for the distributions A and C of Table 15.7 and show in each the relative position of the mean, the median and the true mode. What can you say about these distributions by considering the positions of the three measures ?

[b] Similarly show the relative positions of the three measures in examples 'g' and 'h' in question 8 below.

(3) *Answer briefly the following* :

[i] How does a change in the size of the class interval affect the modal value of a distribution ? Illustrate.

[ii] Suppose that a group of 50 high school boys and 50 high school girls were given a 'rope climbing' test or 'throwing a ball' test. How do you believe the scores would be distributed ? Illustrate by a frequency polygon or a curve.

[iii] The median height of two equal groups of boys is 60" and 64" respectively. Can we conclude that the median height of the total group is 62" ? Give reasons.

[iv] The authorities of a certain school prefer to report the mean salaries of their teachers, while the teachers prefer to report the median salary of their group. Why do they prefer different measures ? Explain this with the help of 'c' in question 8 below.

[v] If a constant is subtracted from each score in a series, what is the effect on the mean ? Utilize this information in finding the mean of the scores : 50, 55, 56, 61, 63.

[vi] Why is the mean called the 'centre of gravity' of the scores ?

[vii] A very easy test is given and many students make perfect or near perfect scores. Illustrate by a frequency curve the distribution and show therein, the position of the mean, the median and the mode.

[viii] If you wish to discount the effect of a few very low scores, which measure of central tendency would you prefer? Why?

[ix] In what respects is the median better than the mean as a measure of central tendency?

[x] What information is revealed by the median but not by the mean?

(4) *Explain the following statements by giving illustrations* :

[i] Mean is sensitive to the size of the extreme measurements.

[ii] Median cannot be called 'the centre of gravity' of the scores.

[iii] Mode is a non-algebraic measure.

[iv] Mode is not a very reliable measure especially with small groups.

[v] It is possible to determine the median graphically from a histogram.

[vi] There is a slight difference in the values of the mean, computed from grouped and ungrouped data.

(5) *Criticize each of the following uses of averages* :

[i] The mean age of men and women at the time of marriage in a certain city based on records for a four year period was 25 years.

[ii] An institution received 1000 donations ranging from Re. 1 to Rs. 150,000, the total being Rs. 260,000. It was announced that the average donation was about Rs. 260.

[iii] A teacher's history class had a mean of 72 and a median of 52 on a standardized achievement test. The teacher concluded that 72 was the average score of her class.

[iv] Two classes have the same mean on a test in arithmetic. Hence, the two groups are identical in arithmetical ability.

(6) *Find the mean and the midscore of the following series* :

[i] 15, 10, 12, 10, 9, 12, 10, 18.

[ii] 47, 51, 38, 28, 58, 74, 54, 42, 38, 63, 65, 84, 86, 49, 81.

[iii] 80, 65, 64, 11, 27, 35, 45, 45, 57, 31, 36, 49, 32, 64, 28, 65.

(7) [a] A test is administered to three classes and the following results are obtained. What is the mean for the whole group ?

Class	Mean	Number of students
A	40	15
B	48	20
C	36	25

[b] The mean monthly pay of 10 workers in a factory is Rs. 110 and the median is Rs. 105. The two best paid workers get an increment of Rs. 15 each. What is the new mean ? The new median ?

[c] Find out the mean of the following scores by the short method : 1002, 10007, 998, 1003, 996, 1000, 1008.

(8) *Compute the mean, the median and the mode for the following distributions :*

[a] Scores	f	[b] Scores	f	[c] Salary in Rupees	f	[d] Scores	f
95–99	4	30	0	300–319	5	93–95	5
90–94	3	29	1	280–299	4	90–92	11
85–89	15	28	2	260–279	2	87–89	12
80–84	6	27	3	240–259	2	84–86	18
75–79	3	26	3	220–239	1	81–83	4
		25	4	200–219	1	78–80	5
70–74	2	24	4			75–77	5
65–69	4	23	1	180–199	4	72–74	7
60–64	5	22	3	160–179	5	69–71	6
55–59	6	21	6	140–159	10	66–68	7
50–54	15	20	2	120–139	8		N = 80
45–49	7	19	1	100–119	12		
	N = 70		N = 30	80– 99	6		
					N = 60		

[e] Scores	f	[f] Scores	f	[g] Scores	f	[h] Scores	f
70–79	3	55–59	2	140–149	8	160–179	1
60–69	6	50–54	7	130–139	14	140–159	2
50–59	0	45–49	10	120–129	15	120–139	1
40–49	0	40–44	15	110–119	20	100–119	3
30–39	4	35–39	18	100–109	15	80– 99	1
20–29	0	30–34	7	90– 99	10	60– 79	2
10–19	3	25–29	4	80– 89	6	40– 59	2
0–9	2	20–24	5	70– 79	5	20– 39	24
N = 18		N = 68		60– 69	3	N = 36	
				50– 59	4		
				N = 100			

(9) Compute the mean, the median and the mode for the six frequency distributions in question 9 given at the end of Chapter 16.

Answers :

(7) [a] 41. [b] mean = 113, median = 105, [c] 1002.

(8)

	a	b	c	d	e	f	g	h
mean	69·5	23·8	169·5	82·0	44·5	39·5	110·0	54·5
median	67·0	24·0	147·5	84·5	49·5	39·5	113·0	34·5
mode	62·0	24·4	103·5	89·5	59·5	39·5	119·0	—

(9)

	a	b	c	d	e	f
mean	13·0	68·25	39·5	66·25	47·0	11·00
median	12·5	67·00	32·0	66·25	48·5	10·75
mode	11·5	66·50	17·0	66·25	51·5	10·25

* * *

16

MEASURES OF VARIABILITY

What is a measure of variability ?

The measures of central tendency are valuable statistical measures, but they describe only one characteristic of the data — the tendency of the scores to pile up at or near the middle of the distribution. Description of test results based wholly on the measures of central tendency is incomplete. A simple example will illustrate this situation. A test in arithmetic is given to two groups; the means of the groups are identical, each being 50

Group	Scores	Mean
A	40, 42, 45, 49, 50, 51, 55, 58, 60	50
B	10, 20, 30, 40, 50, 60, 70, 80, 90	50

Even an inexperienced teacher will recognize that very different ranges of ability are present in the two groups A and B. The first group is decidedly more homogeneous and the second is decidedly more heterogeneous (Fig. 16.1). We may expect the

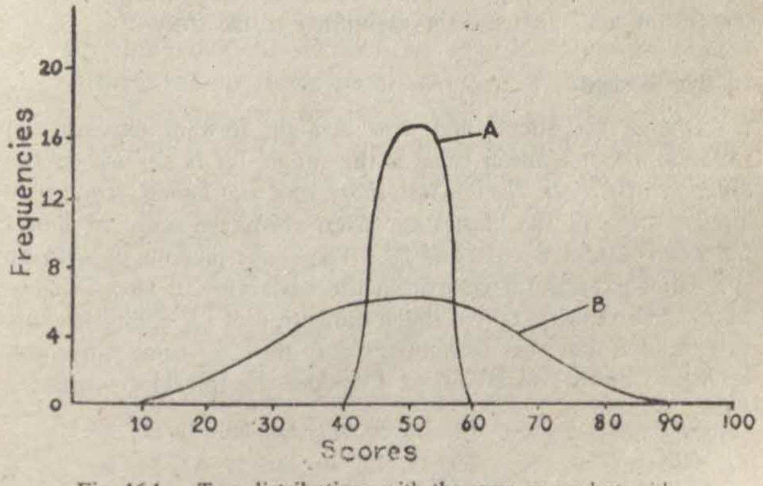


Fig. 16.1 : Two distributions with the same mean but with widely different variabilities.

first group to grasp new ideas and to progress at about the same rate while we should expect the second group to show considerable disparity in the speed of grasping new ideas. Thus the two groups differ from one another in variability, though they have the same mean. Variability is an important feature of a statistical series. The mean, median and mode conceal this feature and a mere comparison with the help of average values may be unfair and misleading.

Fig. 16.1 shows two frequency distributions of the same area and the same mean (50), but of very different variability. Group B ranges from 10 to 90 and group A from 40 to 60. Group B is four times as variable as group A. The flatness or peakedness of the curve is the graphical indication of the variability of the data. Other things being equal, less peaked the curve, greater is the variability. It will thus be clear that we need to take variability into account in describing and comparing statistical data. Variability may reveal more about the series than an average value. *A measure of variability is an indicator showing the extent to which the scores tend to scatter or spread about a measure of central tendency.* The amount or extent of variability of a statistical series is described quantitatively by various measures, the most common of which are the *range*, the *quartile deviation*, the *mean* or *average deviation* and the *standard deviation.* All these measures represent distances rather than points and the larger they are, greater is the variability of the scores.

1. The Range

One of the simple and most straight forward measures of variation about a mean value is the range. *It is defined as the difference between the highest score and the lowest score in a distribution.* In the illustration given above the range of scores for group A and B is 20 and 80. The range presents us with an approximate basis for comparing the variability of two distributions. The range is only a rough indication of variability because it is possible that two distributions may have the same range and yet have different variabilities. Consider the following scores :

Group B	10, 20, 30, 40, 50, 60, 70, 80, 90.
Group C	10, 47, 48, 49, 50, 51, 52, 53, 90.

Here the two groups have the same range, but the variability is different. Scores in distribution B are scattered over the entire range, while most of the scores of group C tend to cluster around the mean. Range takes into account only the two extreme scores and any change in these scores will cause a fluctuation in the range. Hence it is an unreliable measure.

When scores are grouped into a frequency distribution as in Table 16.1, the individual items lose their identity and there is no way of determining the actual range of the series. For grouped data, the difference between the higher real limit of the top class and the lower real limit of the bottom class may be taken as the range. In the distribution in Table 16.1, the range is 139·5 – 59·52 = 80·00. Ordinarily the range is used with reference to ungrouped data.

Limitations of the Range : (i) It is an unreliable measure of variability as it takes into account only the two scores at the two extremeties. (ii) It should not be used when two distributions have a markedly different number of cases. (iii) Similarly it should not be used for comparing two distributions where the units of measurement are not the same, for example, if the heights of two groups are given in inches and in centimeters respectively, the range would be meaningless.

2. The Quartile Deviation (Q)

We have seen above that the range is a function of the scores at either end of the distribution. In order to remove the influence of these two end scores, a special type of range called the *inter-quartile range*, has been developed as a measure of variability. Quartiles are the points which divide a distribution into four equal parts. The first quartile Q_1 is defined as the point on the scale of scores below which 25 percent or a quarter of the scores lie; the second quartile Q_2 or the median is the point below which lie 50 percent of scores and the third quartile Q_3 is the point below which lie 75 percent of the scores. Thus Q_1, Q_2 and Q_3 divide a scale of scores into four intervals, each of which includes 25 percent or one quarter of the scores. Q, the quartile deviation is one half of the distance between the first and the third quartiles. Hence it is often referred to as the *semi-interquartile range*.

$$Q = \frac{Q_3 - Q_1}{2} = \frac{\text{75th percentile} - \text{25th percentile}}{2} \quad \ldots\ldots\ldots (9)$$

Computation of the Quartile Deviation : To compute Q, we must first get the 75th and 25th percentiles. These statistics are found in exactly the same way as was the median which is the 50th percentile or Q_2. The only difference is that $\frac{1}{4}$ of N is counted off from the low end of the distribution to find Q_1 and that $\frac{3}{4}$ of N is counted off to find Q_3.

$$Q_1 = l + i \times \left(\frac{N/4 - F}{f}\right) \quad \ldots\ldots\ldots (10)$$

$$Q_3 = l + i \times \frac{\left(\frac{3N}{4} - F\right)}{f} \quad \ldots\ldots\ldots (11)$$

where

l = exact lower limit of the interval in which the quartile falls,

i = size of the class-interval, N = number of scores,

F = cumulative frequency up to the interval which contains the quartile,

f = frequency on the interval containing the quartile.

Table 16.1 : Computation of the Quartile Deviation

1 Scores X	2 Frequency f	3 Cumulative Frequency	4 Procedure
130–139	2	80	Computing Q_1, the 25th percentile : $\frac{1}{4}N = \frac{80}{4} = 20$ Counting up : 3 + 7 = 10 $\therefore$ Q_1 lies in the class interval 80–89 which has the frequency $f = 16$ Lower limit of the class-interval in which Q_1 falls, $l = 79{\cdot}5$ No. of scores upto the quartile interval, F = 10 $\therefore\ Q_1 = 79{\cdot}5 + 10\left(\frac{20-10}{16}\right)$ $= 79{\cdot}5 + 6{\cdot}25 = 85{\cdot}75$
120–129	7	78	
110–119	11	71	
100–109	14	60	
90–90	20	46	
80–89	16	26	
70–79	7	10	
60–69	3	3	
	N = 80		

Similarly compute Q_3 $\left[\frac{3N}{4} = 60, F = 46, f = 14, l = 99{\cdot}5\right]$

$$Q_3 = l + i \times \frac{(3N/4 - F)}{f} = 99{\cdot}5 + 10\left(\frac{60 - 46}{14}\right)$$

$$= 99{\cdot}5 + 10 = 109{\cdot}5$$

Quartile Deviation or Semi-Interquartile Range

$$Q = \frac{Q_3 - Q_1}{2} = \frac{109{\cdot}50 - 85{\cdot}75}{2} = \frac{23{\cdot}75}{2} \doteq 11{\cdot}875 \ldots (9)$$

The quartile deviation is sometimes spoken of as the average deviation of the first and third quartiles from the median. The value of the median (Q_2), for the above distribution is 96·5 (Table 15.5). Hence the deviation of the quartiles Q_1 and Q_3 from the median (Q_2) are:

$$Q_3 - Q_2 = 109{\cdot}5 - 96{\cdot}5 = 13{\cdot}00$$

$$Q_2 - Q_1 = 96{\cdot}5 - 85{\cdot}75 = 10{\cdot}75$$

$\therefore$ Q = Average of these two deviations = $\frac{13{\cdot}00 + 10{\cdot}75}{2} = 11{\cdot}875$

Check this value with the one given above.

Quartiles should not be confused with quarters. Note that Q_1, Q_2 and Q_3 are points on the measuring scale. They are division points between the scores. A pupil scoring above Q_3 is in the highest quarter but not in the highest quartile, since this expression is meaningless. An individual may have his score *at* the second quartile but not *in* the second quartile. *Quartiles are only points and not ranges, while quarters denote range.*

Merits of Q: (i) Since it is independent of the values of the extreme items in a series, the quartile deviation is a more representative and trustworthy measure of variability than the over-all range. (ii) Since Q measures the average distance of the quartile points from the median (Q_2), it is a good index of the score density at the middle of the distribution. If the scores in the distribution are packed closely together, the quartiles will be near to one another and Q will be small; if they are widely scattered, the quartile points will be relatively far apart and Q will be large (Fig. 16.1). (iii) Quartiles are useful in indicating the swewness of a distribution. If the distribution is symmetrical, Q_1 and Q_3 are equidistant from the median. When there is any skewness in the distribution, the two distances will be unequal.

Skewness is positive when $(Q_3 - Q_2) > (Q_2 - Q_1)$
Skewness is negative when $(Q_3 - Q_2) < (Q_2 - Q_1)$
Skewness is zero when $(Q_3 - Q_2) = (Q_2 - Q_1)$

By using this criterion we find that the distribution in Table 16.1 is positively skewed to a moderate degree. (iv) Like the median, the measure is applicable to open end distributions and also applicable when the units of measurement are unequal.

Limitations : (i) Being based upon the range of the middle 50 percent of the items, it tells nothing about the variability of the lower and upper 25 percent of scores. It is possible for two distributions to have equal Q's, but quite dissimilar variability in the lower and upper quarters. (2) Like the median, it is a non-algebraic measure, and hence quartile deviations of two or more distributions cannot be combined to give the quartile deviation of the combined distribution. (3) Methods of determining the quartile deviation for ungrouped data are not satisfactory, because the points Q_1 and Q_3 can be defined exactly only in a frequency distribution. (4) The quartile measures computed from a series too short to warrant grouping are ordinarily of very limited use.

3. The Average Deviation (AD)

The average deviation also called as the *Mean Deviation* (*MD*) is the mean of the deviations of all the separate scores in a series, taken from their mean. When a score coincides with the mean, its deviation is zero. Deviation is represented by the symbol x.

$$x = X - M \quad \left[\begin{array}{l} X = \text{a score in the series} \\ M = \text{arithmetic mean of the series} \end{array} \right]$$

[When the deviation of a score is taken from the assumed mean, it is denoted by the symbol x'. See Tables 15.4 and 16.6]

If we take the algebraic sum of the deviations of all scores, the sum would be zero. This makes it arithmetically impossible to work with. We can solve this difficulty by disregarding the signs of the deviations. In doing so, we disregard their direction; we are interested only in their magnitudes. The formula for the AD when scores are ungrouped would then be

$$AD = \frac{\Sigma|x|}{N} \quad \ldots\ldots\ldots\ldots\ldots (12)$$

where $|x|$ indicates the absolute value of x, disregarding the algebraic sign and N is the number of scores. Sometimes deviations are taken from the median, instead of from the mean, in which case we specify that the deviation is taken from the median.

(a) *Computation of AD from Ungrouped Data* : Consider the following scores; the mean (M) of the series is 50. Find the deviation of each score from the mean; take the sum of the absolute values of the deviations of the scores and divide it by N.

Table 16.2 : Computation of AD from ungrouped data

Scores (X)	40	42	45	49	50	51	55	58	60
Deviation from the mean (x)	–10	–8	–5	–1	0	1	5	8	10

$$\therefore AD = \frac{\Sigma|x|}{N} = \frac{48}{9} = 5{\cdot}33 \quad \ldots (12)$$

(b) *Computation of AD from grouped data* : When data are grouped, the mid-point of the interval represents every score in the interval. Table 16.3 illustrates the method of computing the AD. The mean of the series is 97·75. Column 2 of Table 16.3 shows the mid-points of the intervals while column 3 shows the frequency. In column 4 are given the deviations of the mid-points from the computed mean. Algebraic signs are recorded for the sake of accuracy but they will not be considered in the computations. Each x deviation in column 4 is then weighted by the frequency which it represents, to give the $f.x$ entries in column 5. The second deviation x of 36·75 is multiplied by f which is 2, the third x of 26·75 by 7 and so on to the end of the column. The sum of the $f.x$ column without regard to signs is then divided by N to give the AD.

$$AD = \frac{\Sigma|f.x|}{N} \quad \ldots\ldots\ldots\ldots (13)$$

Interpretation of AD : In samples that are not too small and when distributions approach the normal curve, the middle 58 percent of the cases should be expected to fall within the limits of 1 AD below the mean and 1 AD above the mean. In Table 16.3, if the distribution of scores is close to normal, we shall expect about 58 percent of students to have their scores between the mean ± 1 AD, that is 97·75 ± 13·49. This would mean that 46 out of 80 scores (58 percent of N) lie between 84·26 and 111·24. The remaining 42 percent of scores are spread beyond these limits.

Table 16.3 : Computation of Average Deviation from grouped data

1 Scores X	2 Mid–point X_m	3 Frequency f	4 Deviation $x = X - M$	5 $f.x$	6 $f.x^2 = (x)\ (fx)$
140–149	144·5	0	46·75	0·00	
130–139	134·5	2	36·75	73·50	
120–129	124·5	7	26·75	187·25	Products in this column may be obtained for computing the standard deviation (S. D.).
110–119	114·5	11	16·75	184·25	
100–109	104·5	14	6·75	94·50	
90– 99	94·5	20	–3·25	–65·00	
80– 89	84·5	16	–13·25	–212·00	
70– 79	74·5	7	–23·25	–162·75	
60– 69	64·5	3	–33·25	–99·75	
50– 59	54·5	0	–43·25	0·00	

$M = 97{\cdot}75 \qquad N = 80 \qquad \Sigma\,|fx| = 1079{\cdot}00$

$$A\,D = \frac{\Sigma\,|f.x|}{N} = \frac{1079}{80} = 13{\cdot}49 \quad \ldots\ldots\ldots \ (13)$$

Merits of AD : (i) It is the simplest measure of variability that takes into account the fluctuations of all items in a series. (ii) It is the most meaningful measure to any one untrained in statistics.

Limitations : (i) Since it is based upon the deviations of all scores, it may be inflated by a single extreme score. When a series is long however and not highly irregular at the extremes, the above limitation is not a serious one. (ii) The measure is unwieldy in mathematical operations. This accounts for the rather infrequent use of this measure.

4. The Standard Deviation (SD)

The standard deviation is the most stable index of variability and is customarily employed in research studies. In computing the AD we disregard the signs and treat all deviations as positive, whereas in finding the SD we avoid the difficulty of signs by squaring the separate deviations. Again the squared deviations used in computing the SD are always taken from the *mean* and never from the *median* or *mode*. It is usually represented by the Greek letter σ called *sigma* and is defined as the square root of the mean of the squares of the deviations of scores from their mean.

(a) *Computation of SD from ungrouped data* : Let us take the same 9 scores which were used in Table 16.2.

Table 16.4 : Computation of SD from ungrouped data

1	2	3	4
Scores X	Deviation $x = X - M$	x^2	Procedure
40	-10	100	(i) Find the deviation of each score from the mean. $x = X - M$
42	- 8	64	
45	- 5	25	(ii) Square the deviations x^2
49	- 1	1	(iii) Sum up the deviations, $\Sigma x^2 = 380$
50	0	0	(iv) Divide the sum by N
51	1	1	(v) Extract the square root of the quotient in (iv) above
55	5	25	
58	8	64	$\text{SD} = \sqrt{\frac{\Sigma x^2}{N}}$ ... (14)
60	10	100	
N = 9 M = 50		$\Sigma x^2 = 380$	$\therefore \text{SD} = \sqrt{\frac{380}{9}} = 6{\cdot}498$

Steps involved in computing the SD are shown in Table 16.4. Computing the SD by this method would be laborious when the value of the mean is not an integral number. In that case, it is better to use the short method, described in Table 16.5 below:

(b) *Computation of SD from ungrouped data by the short method* : It will save time and computation to apply the short method directly to the ungrouped data (Table 16.5). Note that the same 9 scores from Table 16.4 are used. It is not necessary to arrange them in any specific order. The assumed mean (AM) is taken at zero and each score becomes at once a deviation x' from this AM, that is, each score X is unchanged. Correction c is the difference between the actual mean (M) and the assumed mean (O).

$$\therefore \quad c = \text{M} - \text{AM} = \text{M} - \text{O}$$
$$= \text{M}$$

$$\therefore \sigma = \sqrt{\frac{\sum x'^2}{\text{N}} - c^2} = \sqrt{\frac{\sum x}{\text{N}} - \text{M}} \qquad \dots \dots \quad (15)$$

because $x' = \text{X}$, and c = M

Table 16.5 : Computation of SD from ungrouped data by the short method

Scores X	x' (or X)	x'^2 or X^2	Procedure
40	40	1600	$\text{AM} = 0$
58	58	3364	$\text{M} = \frac{\sum \text{X}}{\text{N}} = \frac{450}{9} = 50$
42	42	1764	$c = \text{M} - \text{AM} = \text{M} - 0$
55	55	3025	$= 50$
50	50	2500	$c^2 = 2500$
45	45	2025	$\sigma = \sqrt{\frac{\sum x'^2}{\text{N}} - c^2}$ ...... (15)
51	51	2601	$= \sqrt{\frac{22880}{9} - 2500}$
49	49	2401	$= \sqrt{2542{\cdot}2 - 2500}$
60	60	2600	$= \sqrt{42{\cdot}2}$
450	450	22880	$= 6{\cdot}496$

(c) *Computation of SD from grouped data* : The process is identical with that used for ungrouped items, except that in addition to squaring the deviation x of each mid-point from the mean, we weight each of these squared deviations by the frequency of the corresponding class-interval. This multiplication gives the fx^2 column. In Table 16.3 we can get the values of fx^2 in column 6. by multiplying the entries x and fx in column 4 and 5. Later on the steps as outlined in Table 16.4 are to be followed. The formula for SD when data are grouped into a frequency distribution is

$$\text{SD or } \sigma = \sqrt{\frac{\Sigma fx^2}{N}} \text{ ....(16)}$$

(d) *Computation of SD from grouped data by the short method* : The short method which was employed in the preceding chapter to calculate the mean, will now be extended in order to compute the standard deviation. The initial steps are identical with those employed to compute a mean (Table 15.4). In order to save the labour of computation, we take the deviations of the mid-points from an assumed mean instead of from the true mean. Deviations are taken in units of class-interval. They are shown by the symbol x' in column 3 of Table 16.6. The remaining steps are described in Table 16.6 on page 323.

(e) *Effect of*

(i) *adding a constant to each score*

(ii) *multiplying each score by a constant* :

(i) If each score in a series is increased by the same amount, say 4, the σ is unchanged, but the mean is increased by the amount of the constant added. If a constant is subtracted, the mean will be decreased by that amount, but the variability (σ) is not affected.

(ii) When the original scores are multiplied by a constant say 10, the net effect of this operation is to multiply both the mean and σ by that constant. Table 16.7 shows the effect of these two operations on page 323.

Interpretation of Standard Deviation

The SD gets its most clear cut meaning for one particular type of distributions of scores. This distribution is called the *normal* distribution. For the normal curve, there is an exact mathematical relationship between the standard deviation and the proportion of cases (Fig. 9.1, page 171).

Table 16.6 : Computation of SD from grouped data by the Short Method

1 Scores	2 f	3 x'	4 fx'	5 fx'^2
130–139	2	4	8	32
120–129	7	3	21	63
110–119	11	2	22	44
100–109	14	1	14	14
90—99	20	0	0	0
80—89	16	−1	−16	16
70—79	7	−2	−14	28
60—69	3	−3	−9	27
N=80			$\Sigma fx'$ =+26	$\Sigma fx'^2$ =224

$$\sigma = i \times \sqrt{\frac{\Sigma fx'^2}{N} - c^2} \quad \ldots (17$$

where

σ = standard deviation.

i = class interval

c = correction = $\frac{\Sigma fx'}{N}$

x' = deviation of a score from the assumed mean.

Steps in the Computation

(1) Assume a value for the mean (AM = 94·5)

(2) Call this AM as zero and show it in the x' (deviation) column.

(3) Lay off the deviations from AM by units of intervals in column 3; indicate the appropriate sign.

(4) Multiply the x' (deviation) values by their frequencies and enter the products in fx' column.

(5) Obtain the squared deviations of the scores in column 5 (fx'^2).

(6) Obtain the correction to the assumed mean.

$$c = \frac{\Sigma fx'^2}{N} = \frac{26}{80} = \frac{13}{40}$$

(7) Obtain the sum of the squared deviations of the scores.

$$\Sigma fx'^2 = 224$$

(8) Obtain the mean of the squared deviations,

$$\frac{\Sigma fx'^2}{N} = \frac{224}{80} = \frac{14}{5}$$

(9) Obtain the corrected mean of the squared deviations of the scores.

$$\frac{\Sigma fx'^2}{N} - c^2 = \frac{14}{15} - \left(\frac{13}{40}\right)^2 = \frac{4311}{1600} = 2 \cdot 70$$

(10) Obtain the standard deviation in class interval units.

$$\sqrt{\frac{\Sigma fx'^2}{N} - c^2} = \sqrt{2 \cdot 70} = 1 \cdot 643$$

(11) Obtain the standard deviation in score units

$$SD = i \times \sqrt{\frac{\Sigma fx'^2}{N} - c^2}$$
$$= 10 \times 1 \cdot 643$$
$$= 16 \cdot 43$$

Table 16.7 : Effect of adding a constant and multiplying by a constant.

Original scores X			Original scores X + a constant.			Original scores multiplied by a constant.		
X	$x = X - M$	x^2	X + 4	$x = X - M$	x^2	(X) × (10)	$x = X - M$	x^2
10	4	16	14	4	16	100	40	1600
7	1	1	11	1	1	70	10	100
6	0	0	10	0	0	60	0	0
4	− 2	4	8	− 2	4	40	− 20	400
3	− 3	9	7	− 3	9	30	− 30	900
$\Sigma X = 30$		$\Sigma x^2 = 30$	$\Sigma X = 50$		$\Sigma x^2 = 30$	$\Sigma X = 300$		$\Sigma x^2 = 3000$
$M = \frac{30}{5} = 6$			$M = \frac{50}{5} = 10$			$\sigma = \frac{300}{5} = 60$		
$\sigma = \sqrt{\frac{\Sigma x^2}{N}} = \sqrt{\frac{30}{5}} = 2{\cdot}45$			$\sigma = \sqrt{\frac{30}{5}} = 2 \cdot 45$ (Formula 14, page 319)			$\sigma = \sqrt{\frac{3000}{5}} = 24 \cdot 5$		

In samples that are not too small and when the distributions approach the normal curve, about two-thirds of the cases should be expected to fall within + 1 and − 1 standard deviation from the mean. If the distribution is exactly normal, 68·26 percent of the scores will fall within the distance of $\pm 1\ \sigma$ from the mean. In the example in Table 16.6, if the distribution is close to normal, we should expect about 68% of the students to have their scores between the mean and ± 1 SD i.e., 97·75 ± 16·43. This would mean that 54 of the 80 scores (68% of 80) lie between 81·32 and 114·18. The remaining 32% of scores are spread beyond these limits.

Rough Checks : In a normal distribution the value of SD is always larger than the value of AD which in turn is larger than the value of Q. These relationships supply a rough check upon the accuracy of the measures of variability. Check the values of these measures as are obtained in Tables 16.1, 16.3 and 16.6. Another rough check is to compare the standard deviation with the total range of scores. In large samples ($N = 500$), the total range is about 6 times the SD; if $N = 100$ it is about 5 times the SD; if $N = 50$ it is about 4·5 times and if $N = 20$ the range is about 3·7 times the SD. Verify from Table 16.6, where the range is 80.

Characteristics of SD : (i) It is taken only from the mean and not from any other measure of central tendency. (ii) If a constant number is added to each of the given scores, the variability is not affected, but the mean of the scores is increased by the amount of the constant added. (iii) When the original scores are multiplied by a constant, the net effect of this operation is to multiply both the mean and SD by that constant.

Merits of SD : (i) It is the most important measure of variability of test scores. It has become a yardstick in terms of which different groups may be compared or the status of a given individual may be evaluated. (ii) SD is also useful in comparing an individual's score on two or more tests, by converting the raw scores into *standard scores*. (Please refer to chapter 18 for a discussion of standard scores.)

Limitations : (i) It is extremely difficult to interpret this measure to those untrained in statistics. (ii) It is more sensitive to the extreme values in a series. By studying two or three series, having a difference in the extreme scores, the student can satisfy himself that the SD is more affected by the extreme values than the two other measures, namely the Q and the AD. (See question 4 at the end of this chapter.)

When will you use the various measures of variability ?

Several considerations come into picture when we decide what measure of variability to employ in any situation. As far as reliability or consistency of the measure is concerned, the most reliable measure would be the SD, followed by AD, Q and the Range. So far as quickness and ease of computation are concerned, the four are almost in reverse order to that just given. If further statistical computation is to be given then the standard deviation is the only measure to be used.

Use the Range : (i) When the quickest possible index of dispersion is wanted. (ii) When the knowledge of the extreme scores is all that is wanted.

Use the Q : (i) When the median is the measure of the central tendency reported. (ii) When the distribution is truncated or incomplete at either end. (iii) When there are a few very extreme scores that would influence the other measures disproportionately. (iv) When the degree of concentration around the median — the middle 50 percent of cases — is of primary interest.

Use the AD : (i) When it is desired to weigh all deviations from the mean according to their size. (ii) When extreme deviations would influence the SD unduly.

Use the SD : (i) When the measure having the greatest reliability is sought. (ii) When further computations that depend upon the SD are likely to be needed. (iii) When interpretations related to the normal probability curve are desired.

FOR FURTHER STUDY

(1) Consider the following scores :

62, 60, 66, 78, 75, 72, 67, 80

[a) Find the mean and σ.

[b] Add 4 to each score and recalculate M and σ.

[c] Subtract 20 from each score and calculate M and σ

[d] Multiply each score by 5 and compute M and σ.

(2) Following results were obtained when a test was given to two sections of grade VIII.

	Class A	Class B
Median	64·6	64·3
Mean	65·0	63·2
75th Percentile	69·0	70·0
25th Percentile	61·0	54·0
S. D.	5·5	10·5

(i) In what respects do the two sections differ ? (ii) What implications do the data have for teaching the two groups ? (iii) How do the two distributions of scores differ in skewness ? (iv) Draw rough curves to show the skewness of the two distributions and show there in the positions of the mean and the median. (v) How do the quartile points help in determining the skewness ?

(3) A test in science is given to a large group of pre-university class students. The mean score is 60 and S.D. is 12. If the distribution of scores is normal, what inferences can you draw about the standing of the following students ?

Jayant	72,	Meena	60,	Suresh	36
Lata	66,	Tanuja	48,	Aruna	84.

(4) *Explain the following statements* :

[i] Averages are more meaningful when accompanied by a statement regarding variability of data.

[ii] Two distributions have equal quartile deviations, yet their standard deviations are found to be different.

[iii] Range is not a reliable measure of variability.

[iv] The standard deviation gets its most clearcut meaning in a normal distribution.

[v] Raw scores of a student on different tests cannot be directly compared.

(5) *Answer the following questions* :

[i] How do the quartile points help in determining the skewness of a distribution ?

[ii] You want to compare the achievement of two classes of the same grade. Which of the statistics will be useful ? Illustrate.

[iii] Explain in statistical terms the difference between a homogeneous and a heterogeneous class. How will you show this difference graphically ?

[iv] Is it possible that a score in a distribution could lie at a distance greater than 2Q from the median ? Explain.

[v] If you wish to discount the effect of a few very extreme scores in a distribution, which measure of variability would you compute ? Why ?

[vi] A teacher's analysis of reading scores in her class indicated that $Q_1 = 26$ and $Q_3 = 46$. She concluded that (a) the median was 36 and (b) that Q was 10. Do you agree with these conclusions ? Explain.

[vii] Two groups of students are given a reading test. For group A the median is 32, and for group B the median is 46. If Jayant has a percentile rank of 62 in group A, can it be concluded that he is equal in reading ability to Madhuri who has a percentile rank of 62 in group B ? Why ?

[viii] Describe a set of data for which either the SD, AD or Q may be most useful.

(6) *Comment on the following statements* :

[i] Quartile deviation for two groups is 10·6 and 14·0 respectively. Hence the quartile deviation for the combined groups would be the average of the two values.

[ii] If the five highest scores in a distribution of 100 scores are increased by five percent, the original values of Q and S.D fo the distribution will be affected.

[iii] The variability of a distribution reveals more about a group than the average value.

[iv] Quartiles are only points and not ranges on a scale.

[v] The SD of a large, approximately normal distribution as computed by a student was 2·0. The over-all range of the distribution was 40.

[vi] Ajit's raw score on two tests is 50 and 60 respectively. Hence his performance on the second test is better than that on the first.

[vii] Standard deviation helps us to interpret an individual's score in a given series.

(7) In two factories A and B, the distributions of salaries are as follows :

	Q_1	Q_2	Q_3
(A)	Rs. 150	Rs. 200	Rs. 230
(B)	Rs. 140	Rs. 200	Rs. 300

[i] What inferences can you draw about these distributions?
[ii] What is the value of Q in the two distributions?
[iii] What do you know about the skewness of the two?
[iv] Which group will have a higher S.D.? Explain.
[v] Do you agree with the view that since the median is the same, the mean salary of the two groups will be the same. Give reasons.

(8) Find the average deviation and the standard deviation of the two series below. Which measure is more affected by the extreme score in the second series?

A : 30, 30, 32, 33, 34, 35, 35, 35, 36, 40
B : 30, 30, 32, 33, 34, 35, 35. 35, 36, 60

(9) *Compute the Q, AD and SD for each of the following distributions* :

[a] Scores	*f*	[b] Scores	*f*	[c] Scores	*f*
27–29	5	90–94	4	90–99	1
24–26	4	85–89	2	80–89	1
21–23	2	80–84	2	70–79	4
18–20	2	75–79	8	60–69	6
15–17	10	70–74	7	50–59	4
12–14	12	65–69	10	40–49	2
9–11	6	60–64	9	30–39	4
6– 8	11	55–59	9	20–29	6
3– 6	7	50–54	4	10–19	14
0– 2	3	45–49	1		N...42
	N...62	40–44	0		
			N...56		

[d] Scores	f
70	6
69	8
68	4
67	4
66	8
65	4
64	4
63	5
62	4
61	1
	N = 48

[e] Scores	f
80–89	5
70–79	10
60–69	15
50–59	22
40–49	20
30–39	12
20–29	13
10–19	7
0–9	4
	N = 108

[f] Scores	f
16–17	5
14–15	4
12–13	4
10–11	8
8–9	5
6–7	4
4–5	2
	N = 32

(10) Compute the Q, AD and S.D. for the eight frequency distributions, given at the end of Chapter 15, question 8.

(11) Following are the scores of two groups on a test in arithmetic. Compare the two distributions. Which statistics will you use to compare the two groups ?

Scores	A	B
80–89	0	5
70–79	1	6
60–69	5	5
50–59	10	7
40–49	20	7
30–39	15	9
20–29	7	10
10–19	2	6
0–9	0	5
	N = 60	N = 60

[i] Compute the value of Q.D.

[ii] Determine the quartile points Q_1, Q_2 and Q_3.

[iii] What is the value of the mean for each distribution? Which aspect of the distributions is not revealed by this value of the mean?

Answers :

(1)

	(a)	(b)	(c)	(d)
M	70	74	50	350
σ	6·91	6·91	6·91	34·55

(2) Class A : positive skewness. Class B : negative skewness.

(6) (i) False (ii) Value of S.D. will change; Q will be the same (iii) True (iv) True (v) Incorrect (vi) Insufficient data (vii) True

(8)

	(A)	(B)
A.D.	2·2	4·8
S.D.	2·8	8·2

(9)

	a	b	c	d	e	f
Q	4·875	8·125	22·5	2·375	14·75	2·85
A. D.	5·9	9·15	21·9	2·27	16·48	2·97
S.D.	7·5	11·25	24·3	2·65	20·10	3·57

(10)

	a	b	c	d	e	f	g	h
Q	16·5	2·375	42·5	7·0	24·6	5·36	15·5	21·25
A. D.	15·5	2·36	57·33	7·125	23·33	6·9	18·53	33·89
S.D.	16·75	2·72	70·00	8·25	25·1	8·6	23·04	41·73

(11) (i) Q: A = 8·5, B = 19

(ii)	Q_1	Q_2	Q_3
A	33·5	39·5	59·5
B	23·5	39·5	61·5

(iii)	mean	Q	AD	SD
[a]	42·5	8·5	10·07	13·0
[b]	42·5	19·0	20·33	23·9

* * *

17

MEASURES OF RELATIONSHIP

In previous chapters we have been concerned primarily with methods of computing statistical measures designed to represent in a reliable way the performance of an individual or a group in some defined trait. Frequently however, it is of more importance to examine the relationship of one variable with another than to measure performance on either alone. No single statistical procedure has opened up so many new avenues of discovery in psychology and education as that of correlation. It is a method of summarizing the relationship between two sets of data. *A coefficient of correlation is a single number that tells us to what extent two things are related, to what extent variations in the one, go with the variations in the other.*

How will you express the relationship between two measures? : Correlation is the most common method of reporting the answer to such questions as : Are certain abilities closely related and others relatively independent? Does a given aptitude test predict success on the job? Are football players good in academic achievement? Is it true that bright children are physically weak than average children? In order to answer such questions we must have a pair of measures for each individual. We can never compute a coefficient of correlation on one person alone, nor can we compute it without having made two sets of measurements on the same individuals or on two matched groups. The method that we use to define this relationship is called correlation. An example might clarify the meaning of correlation.

Individual	Test X	Test Y
A	20	52
B	17	47
C	15	39
D	13	21
E	12	15

Individual	Test M	Test N
A	17	20
B	15	32
C	14	58
D	13	84
E	12	92

Notice that the relationship between tests X and Y is perfect. Individual A scored the highest on both tests. Individual E scored the lowest. When we assign the rank for each score for every individual, we see that they agree perfectly on both tests. This is *perfect positive correlation;* this relationship is denoted by the correlation coefficient of + 1·00. We may also have the reverse situation. Consider the tests M and N. There is also a relationship between the two tests of scores. Individual A scored highest on test M but lowest on test N. The ranks for the other individuals are similarly reversed. This is still *perfect correlation*, but in a *negative direction*. This is denoted by the correlation coefficient of – 1·00.

It is also possible that individual A might rank first on test I, but third on test II, individual C might rank third on test I but fifth on test II and so on. In other words there would be no pattern of relationship shown in the data. The correlation coefficient would be zero, that is, there is no relationship.

Degrees of Correlation

The correlation coefficient can vary from perfect positive + 1·0 to perfect negative – 1·0. The extreme values are rarely obtained in practice but high values such as + ·89 or + ·90 are common. As the relationship ranges from 0·0 to 1·0, the relationship becomes greater until we reach + 1·0, where the relationship is perfect. As the coefficient ranges from 0 to – 1·0, the same holds true, the correlation becomes greater but in a negative direction. The area of a rectangle is doubled if its length is doubled; the area becomes one half, if the side of the rectangle is reduced to one half. As the temperature increases, the height of the mercury column in a thermometer also increases by the same extent. These are examples of perfect positive correlation. The two variables vary together in the same direction. It is also possible that the variables may vary exactly in opposite directions, for example, greater the speed of a car, less is the time required to cover a given distance.

In the natural sciences correlations between many phenomena tend to be perfect or nearly so. A number of examples showing perfect relationship have been given above. But in the biological and social sciences, correlations are much less marked. They rarely show perfect relationship. For example, height and weight vary together though not perfectly. The correlation is weakened, by cases such as 'short and heavy' or 'tall and light' men. Similarly the relationship between intelligence test scores and school marks although significant, is not perfect. Husbands and wives tend to be more like each other with respect to age, amount of education and many other factors than they are like other people in general. The sons of tall fathers tend to be taller than average and the sons of short fathers tend to be short. These are all examples of *positive imperfect correlation.*

Negative imperfect relationship is found between (*i*) self-confidence and submissiveness, (*ii*) intelligence and suggestibility, (*iii*) school achievement and absence from school, (*iv*) speed and accuracy of doing a work and a number of such other traits. Near zero correlation might be expected between (*i*) speed of tapping and reasoning ability, (*ii*) ability to match colours and skill in archery or (*iii*) shape of head and intelligence. These latter variables do not vary together. You cannot judge intelligence by measuring the shape of head of various individuals. (For additional examples of correlation refer to Table 17.3 at the end of this chapter.)

How to compute Coefficient of Correlation

There are several methods of finding the coefficient of correlation between two series of measures. If the data are in the form of ranks rather than raw scores, we can apply Spearman's *Rank-Difference Method.* The coefficient is indicated by the lower case Greek letter *rho* (ρ). If the data are in the form of raw scores, it is necessary to translate them into rank order. The procedure is explained in Table 17.1. There we have ten pairs of scores for ten individuals on two variables namely, suggesibility and intelligence. The problem is to find the relationship between these two variables.

Following are the steps in computing ρ *by the Rank-Difference Method* : (1) Rank the individuals from the highest to the lowest in the first variable, (here it is the suggestibility score) and call these as ranks R_1 (Column 4). The highest score receives rank 1, the next highest, 2, etc. The only difficulty encountered is when we find ties. Thus the two individuals C and H have the same score of 52, and each is ranked 4·5. Instead of ranking one man 4 and the other 5, we simply average the ranks and give each man a rank of 4·5. The next rank is 6. Similarly three individuals B, E and G have a score of 37. We treat them alike and assign a rank of 8, by averaging the three ranks 7, 8 and 9. If the lowest person is not tied with others, the last rank should be equal to N.

(2) Similarly rank the second list of scores and call them R_2 as shown in column 5. Here the lowest two persons A and J are tied together and hence they are assigned the rank 9·5.

(3) In column (6), the difference (D) between each pair of ranks is entered. The sum of D's is zero. *Do not proceed further if you do not get the algebraic sum of column 6 as zero.* This serves as a check to the accuracy of ranking.

(4) Square each difference to find D^2 (column 7). Sum up the squares to find ΣD^2. Here it equals 269·000.

Table 17.1 : Computation of ρ by Spearman's method of rank-difference.

1 Subject	2 Suggestibility score X	3 IQ Y	4 R_1	5 R_2	6 $D=R_1-R_2$	7 D^2
A	64	95	2	9·5	−7·5	56·25
B	37	105	8	4	4·0	16·00
C	52	109	4·5	2	2·5	6·25
D	53	98	3	8	−5·0	25·00
E	37	99	8	7	1·0	1·00
F	40	102	6	6	0·0	0·00
G	37	103	8	5	3·0	9·00
H	52	108	4·5	3	1·5	2·25
I	28	114	10	1	9·0	81·00
J	70	96	1	9·5	−8·5	72·25

N = 10 $\Sigma D = 0$, $\Sigma D^2 = 269·00$

$$\rho = 1 - \frac{6\,\Sigma D^2}{N(N^2-1)} = 1 - \frac{6 \times 269}{10(100-1)} = -\,0·63$$

(5) Compute the coefficient of correlation ρ by the formula,

$$\rho = 1 - \frac{6 \Sigma D^2}{N(N^2-1)} \quad \ldots \ (18$$

(rank difference correlation coefficient)

Where ρ = rank difference correlation coefficient

ΣD^2 = sum of the squares of difference in rank

N = number of pairs

Interpretation : There is a substantial relationship between intelligence test score and suggestibility score, but the relationship is in inverse direction. It indicates that higher the intelligence, less is the suggestibility of the individual. Students scoring high on one test, would generally score low on the other test.

It will be observed from the above formula that the size of the rank differences directly affects the size of the coefficient. If the correlation is perfect, that is, if the ranks are the same for both tests, each rank difference in the D column will equal zero and the corresponding D^2 column will be zero. As a result, the correlation is + 1·0. As the relationship between the scores becomes poorer, the rank differences increase and the fraction to be subtracted from 1 (see the formula) becomes larger. In the case where there is a negative correlation, rank differences are very large and the fraction is greater than one, which obviously results in a negative coefficient.

Sometimes instead of giving raw scores, ranks obtained by students in the particular subject or skill will be given. Then you can drop the column R_1 or R_2 (Table 17.1) and proceed directly to find the difference in ranks. See examples 10 (i), (iii) and 11 (ii) and 12 at the end of this chapter.

Another widely used method is the one devised by the English statistician Karl Pearson about the year 1900. This method does not use the ranks of the scores but utilizes the actual sizes of the scores in the computation. This is called the Pearson *product-moment method* and the coefficient is called the Pearson *r*.

In the various examples quoted in the earlier section, the relationship can be expressed in the form of a graph. When the graph is a straight line, the relationship is said to be linear. This happens when an increase in the value of one variable is accompanied *consistently* either by an increase or a decrease in the value of the other variable. Sometimes however, the direction of relationship may differ at different levels of one of the variables. Physical strength tends to increase with chronological age up to early adulthood and then decreases with the approach of old age. The length of a shadow of an object decreases as the time passes during the first half of the day and increases during the latter half of the day. If this relationship is plotted, we get a curved line. Such combinations of both positive and negative relationships are called *curvilinear*. Special methods are required for determining the degree of curvilinear relationship. These methods are beyond the scope of this book.

Interpreting the Coefficient of Corrleation

In interpreting this coefficient, two things must be considered. The first is the sign of the coefficient. The sign indicates the direction of relationship. Positive coefficients indicate direct relationship, that is, there is a tendency for the two series to vary in the same direction while negative coefficients indicate inverse relationship, that is, the two variables have a tendency to vary in opposite directions. An r of $+ 0{\cdot}82$ indicates direct covariation while an r of $- 0{\cdot}82$ denotes inverse relationship. But both indicate a high relationship.

Another thing is equally important and far more difficult to interpret; it is the magnitude or size of the coefficient which indicates the degree or closeness of the relationship. The minimum coefficient is zero, which indicates no consistent relationship whatsoever. From this minimum value the coefficients increase in both directions until $- 1{\cdot}00$ is reached for one limit and $+ 1{\cdot}00$ for the other. It should be noted that both $- 1{\cdot}00$ and $+ 1{\cdot}00$ indicate equally close relationship, for both are perfect. In like manner all other values of the same size such as $- 0{\cdot}71$ and $+ 0{\cdot}71$ indicate equally close relationship. It is the size not the sign of the coefficient that gives the clue to the closeness or degree of relationship. Classification given in Table 17.2, will be useful in interpreting a given r.

Table 17.2 : Interpreting the coefficient of correlation.

(1) r from 0·0 to ± 0·2	denotes slight correlation; indifferent or negligible relationship
(2) r from ± ·20 to ± ·40	denotes low correlation; definite but small relationship
(3) r from ± ·40 to ± ·70	denotes moderate correlation; substantial relationship
(4) r from ± ·70 to ± ·90	denotes high correlation; marked relationship
(5) r from ± ·90 to ± ·99	denotes very high correlation; very dependable relationship
(6) r of ± 1·00	denotes perfect relationship

This classification is broad and somewhat tentative and can only be accepted as a general guide with certain reservations. The coefficient of correlation should be judged with regard to the purpose for which it is computed. Thus an r or + 0·65 between an intelligence test and achievement in school may be considered as fairly high while an r or + 0·65 between scores on two parallel forms of the same test may be considered as low.

Precautions in interpleting the Correlation Coefficient : (1) The coefficient is a complex function of the degree of relationship and two coefficients cannot be directly compared. A given r cannot be interpreted directly as a percentage. An r of 0·00 represents no relationship at all, but an r of 0·68 does not mean 68 percent relationship. Similarly we cannot say that an r of 0·50 indicates two times the relationship that is indicated by an r of 0·25. Nor can we say that an increase in correlation from r = ·40 to r = ·60 is equivalent to an increase in r from ·70 to ·90.

(2) The second precaution is that correlation does not necessarily mean causation. A high correlation is often erroneously interpreted as showing that one factor causes the other. For example, suppose there is a high correlation between scores on a vocabulary test and scores on a reading test. Does it mean that good vocabulary causes one to be a good reader ? Possibly yes. Or does ability to read well cause one to acquire a good vocabulary ? An equally likely explanation. But it is also likely that some other variable such as high intelligence or favourable

environment or superior teaching methods is the cause of good vocabulary and also of good reading ability. Thus there might be a common cause which serves to explain the relationship. Sometimes it is quite possible that the correlation found between two variables is entirely coincidental, for example, there might be a positive correlation between the number of persons taking alcoholic drinks and the number of students attending the colleges. Similarly the salaries of teachers and the price of milk over a period of years may show high correlation. But we cannot say that one variable has caused a change in the other. It might be purely coincidental. Note carefully that while correlation does not directly establish a *causal relationship,* it may furnish clues to causes and these can be taken advantage of in further studies.

Uses of Correlation Coefficient

(1) It is used in determining the *validity* of a test. In validating a group intelligence test for example, we find the correlation between scores on the group test and scores on an individual test of intelligence. A high correlation would indicate that the two tests go together. (2) It is used in determining the *reliability* of a test. The test is given to the students on two occasions and correlation between the scores of each individual on the two occasions is found out. This helps us to know how precise and consistent the test is. (3) If answerbooks have been assessed by two examiners, the *objectivity* of their scoring can be determined by finding the coefficient of correlation between the marks given by the two examiners. [See question 4 at the end of this chapter.] (4) Sometimes we can test the validity of arguments put forth by people such as for example, intelligence depends upon the size of the head or artistic ability depends upon the length of the fingers etc. Coefficients of zero or near zero will discredit the superstitious belief regarding the causal relationship between the size of head and intelligence. (5) Is the preliminary examination useful in predicting the success or failure of pupils at the S. S. C. Examination ? We can answer such questions by computing the r between the scores on the preliminary and on the final examination. The size of r will also help us in determining the need of the preliminary examination. (6) Whenever the question of deciding the relationship between two variables arises, we can use this coefficient to weigh the arguments.

Table 17.3 : Correlation Coefficients of Selected Variables

Variables	*r*	Classification
1. Side of a square versus its diagonal	1·00	Positive perfect
2. Quantity of rice and its price	1·00	-do-
3. Height of identical twins	0·96	Very high
4. IQ on the two forms of 1937 Stanford Binet	0·93	-do-
5. Intelligence of identical twins	0·88	High
6. Height versus weight	0·60	Moderate
7. Intelligence of siblings	0·53	-do-
8. Height of siblings	0·50	-do-
9. Vocabulary and spelling (high school students)	0·50	-do-
10. English and mathematics (college students)	0·49	-do-
11. Mathematics and history (college students)	0·44	-do-
12. English and drawing (college students)	0·25	Low correlation
13. Grip of the hand and tapping	0·25	-do-
14. Strength of the grip and speed of running	0·16	Slight
15. English and art (college students)	0·15	-do-
16. Height versus Binet IQ	0·06	Near zero
17. Shape of head versus intelligence	0·01	-do-
18. Height versus sociability	0·00	-do-
19. Number of physical defects among boys versus school progress	– 0·29	Low
20. Absence from school versus school achievement	– 0·40	Moderate
21. Rate of road accidents and safety devises	– 0·48	-do-
22. Memory test versus age of adults from 20 to 90	– 0·50	-do-
23. Price of a commodity and its availability	– 0·70	High
24. Speed of a car versus time required to cover a certain distance	– 1·00	Perfect negative
25. Volume of a gas and its pressure	– 1·00	Perfect -do-

A few correlation coefficient : Table 17.3 gives a number of different correlations that have been reported for different types of variables. It should be noted that some of these coefficients may slightly change, depending upon the sample and the test used. Perfect correlations will not however change with the samples. It will be seen from Table 17.3 that in general, abilities are positively rather than negatively interrelated though the relationship is usually moderate than close.

FOR FURTHER STUDY

(1) What correlation would you anticipate in each of the following pairs of variables? How will you interpret the same?

[i] Achievement in reading and achievement in spelling.
[ii] Life of automobile tyres and speed of driving.
[iii] Direction of the wind and the movement of smoke from a chimney.
[iv] Age in January and age in December the same year and of the same individuals.
[v] Salaries and years of experience of teachers.
[vi] Amount of unemployment and retail sales.
[vii] Size of head and intelligence.
[viii] Husbands' and wives' height.
[ix] Speed of sound and temperature.
[x] Cost of production and the selling price of an article.
[xi] Chronological age and mental development.
[xii] Conservative attitudes and education.
[xiii] Size of vocabulary and ability to learn.
[xiv] Speed and accuracy of doing a complicated task.
[xv] Intelligence and scholastic achievement.
[xvi] Diameter of a circle and its perimeter.

(2) Is there a causal relationship between each pair of variables in question 1 above? Comment on the possible causal relationships.

(3) Name two pairs of each type, not listed above, which would show (i) perfect positive correlation, (ii) imperfect positive, (iii) zero correlation, (iv) imperfect negative and (v) perfect negative correlation.

(4) Explain the term 'correlation'. How is correlation expressed? What type of questions can be answered by using the correlation technique?

(5) *Answer the following questions* :

[i] What would be the size of r, if we can predict perfectly performance on one test from that on the other ? Illustrate.

[ii] Why is it difficult to interpret a given value of r such as ·55 or ·40 ?

[iii] What are the uses of the coefficient of correlation ?

[iv] What factors will you consider while interpreting a given coefficient of correlation ?

[v] What will be the effect on the value of r if all the scores on the variable X or Y are decreased by a constant ? (Illustrate by considering the problem in Table 17.1).

[vi] Describe a specific situation in which an r of zero or near zero would be important.

[vii] The principal of a college says, " Marks obtained by students in English at the S.S.C. examination are the best single predictor of success at the first year college examination. " How can you prove or disprove this assertion ? Illustrate.

[viii] It is frequently said that, unless pupils can read well, they do badly in other school subjects. How would you find out in which subjects this state of affairs tends to be most marked ?

[ix] Which of the following r's has the least predictive value ? Why ? ·91, ·50, ·17, $-$·25, $-$1·00

[x] What rank would be assigned to a score of 97 and 95, in the following distribution, if the rank of the lowest score 93, is 9?

96, 97, 93, 94, 97, 95, 95, 94, 95.

[xi] What precautions should be taken in interpreting a given correlation coefficient ?

(6) *Explain the following statements with proper illustrations* :

[a] In interpreting a given coefficient of correlation both the sign and size of the coefficient are important.

[b] The fact of correlation does not demonstrate sequence.

[c] A negative coefficient of correlation indicates inverse relationship.

[d] The size of the coefficient of correlation indicates the degree or closeness of relationship.

[e] An r of 0·00 shows no consistent relationship whatsoever between two variables.

[f] Correlation does not necessarily mean causal relationship.

(7) *Comment on the following statements* :

[i] A coefficient of + ·90 shows closer relationship than that shown by an r of – ·90.

[ii] An r of ·40 indicates two times the relationship that is indicated by an r of ·20.

[iii] The minimum value of r is 0·00 and not – 1·00.

[iv] A given value of r such as ·58. may be considered 'high' in one situation but 'low' in another situation.

[v] Two values of r cannot be directly compared.

[vi] The terms 'high' and 'low' used to indicate the size of an r are misleading, if taken at their face value.

[vii] An r of 0·00 does not provide us any information.

[viii] A student computed a coefficient of correlation between two scores in English and French and found it to be 1·40.

[ix] An r of ·65 shows high relationship between two variables.

[x] A teacher obtained an r of ·92 between a reading test and a general knowledge test. She concluded that a good score on the general knowledge test, is the result of good reading ability.

[xi] Negative correlations may be disregarded, since they have no practical value.

(8) *Explain the meaning of each of the following correlation coefficients*

[i] r between the scores on a reading test and an intelligence test is + 0·82.

[ii] r between the speed of running and achievement in English is + 0·03.

[iii] Ratings of pupils on 'good citizenship' and on 'aggressiveness' show a correlation of – 0·56.

[iv] r between a teacher's opinion about pupil's progress and marks in the final examination is + 0·36.

[v] r between two parallel forms of a test is + 0·92.

(9) What do you understand by the term correlation ? How is the correlation between two variables determined ? What are the uses of a correlation coefficient ?

(10) Compute by the rank difference method the coefficient of correlation between the two sets of scores in the four examples given below. Interpret the result in each case. What precautions should be taken while interpreting a given coefficient of correlation ?

	(i)		(ii)		(iii)		(iv)	
Subject	Marks in English	Rank order in a mathematics test	Marks in Algebra	Marks in Geometry	Rank in debate	Rank in drama activity	Marks in debating	Marks in athletics
A	45	1	77	54	8	1	44	60
B	53	5	65	22	10	8	47	65
C	68	2	54	63	5	10	53	66
D	38	8	43	70	7	5	48	76
E	25	6	32	35	1	6	45	52
F	39	7	31	12	3	7	44	70
G	49	4	29	21	4	2	41	66
H	56	9	27	9	2	3	50	56
I	31	10	11	8	6	4	56	62
J	72	3	4	6	9	6	41	49

(11) What are the uses of the coefficient of correlation ? Compute the coefficient of correlation by the rank difference method between the sets of marks in each of the following four examples. Interpret the result in each.

[In example 11 (a), student E will have the first rank and students I, H and F will follow in order, in the ranks of the first variable.]

	(a)		(b)		(c)		(d)	
Student	No. of errors in an objective test	Marks obtained in the test	Years of teaching experience	Rank in teaching efficiency	Percentage at the preliminary examination	Percentage of marks at the S.S.C. examination	Scores in a science test	Scores in a history test
A	1	83	10	10	70	60	65	70
B	4	75	5	2	75	74	41	64
C	0	87	6	4	77	75	66	64
D	3	80	16	1	57	50	38	72
E	17	56	7	8	57	60	45	51
F	9	62	12	3	90	80	65	82
G	7	66	2	9	95	65	44	75
H	11	61	3	12	66	42	58	62
I	12	60	10	7	90	72	65	84
J	6	68	5	11	57	63	42	64
K	8	65	4	5	60	30	49	80
L	5	70	1	6	65	38	38	55

(12) Ten students had written an essay which was examined by three examiners, Leela, Vijaya and Judy. Ranks as given by the three examiners are given below.

By using the rank difference method, discuss which pair of examiners have a common approach of examining the essays.

(Hint : Calculate the coefficient of correlation between the rankings of : (i) the first and the second examiner, (ii) second and third, and (iii) the third and the first examiner. The highest positive coefficient will indicate similarity of approach. Since ranks are already given, you can directly find the difference between the ranks as given by two examiners.)

Student	First examiner (a)	Second examiner (b)	Third examiner (c)
A	1	3	6
B	6	5	4
C	5	8	9
D	10	4	8
E	3	7	1
F	2	10	2
G	4	2	3
H	9	1	10
I	7	6	5
J	8	9	7

Answers

10. (i) $\rho = +0{\cdot}49$ (ii) $+0{\cdot}84$ (iii) $+0{\cdot}20$ (iv) $+0{\cdot}19$

11. (a) $\rho = -1{\cdot}00$ (b) $+0{\cdot}4$ (c) $+0{\cdot}65$ (d) $+0{\cdot}35$

12. ρ between a and b $= -0{\cdot}2$, between b and c $= -0{\cdot}3$, between c and a $= +0{\cdot}64$. The first and third examiners have a common approach in assessing the essays. The other two pairs have divergent views.

18

THE NORMAL CURVE

The Meaning of Normal Distribution :

In figures 18.1 and 18.2 are a histogram and a polygon which represent frequency distributions of data from psychology and physiology. It is apparent, even upon superficial examination, that both the graphs have the same general form, the measures are concentrated closely around the centre and taper off from this central high point or crest to the left and right. There are relatively few measures at the 'low-score' end of the scale, an increasing number up to a maximum at the middle position, and a progressive tailing off toward the 'high-score' end of the scale. All the measures moreover, are symmetrically distributed; that is to say, there are as many measures on one side of the distribution as on the other side.

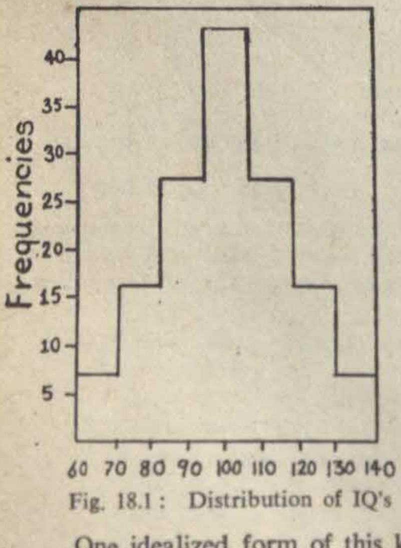


Fig. 18.1 : Distribution of IQ's

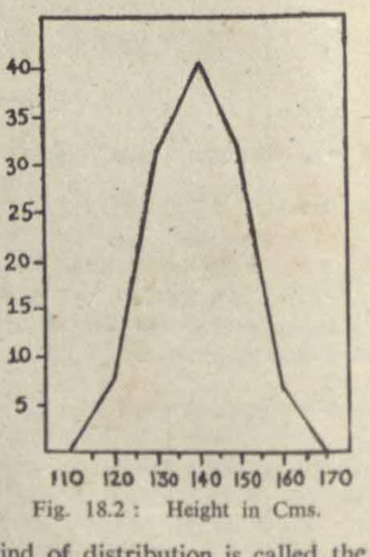


Fig. 18.2 : Height in Cms.

One idealized form of this kind of distribution is called the *normal probability curve* (Fig. 18.3) and is typical of many biological, psychological, and physical measurements. It occurs, in fact, wherever there is a large element of chance entering into a measurement and when each chance event has about a 50–50 probability of occurring. One of the best ways, for example, to obtain such a distribution is to toss coins. You could put 10 coins in a jar, shake them up, toss them on the table and count the number of heads from 0 to 10 that appear on each toss. The measurement in this case is the number of heads that turn up. If you repeat this time after time, for say, 100 tosses, and each time tabulate the number of heads, the distribution you get approaches the *normal probability curve* such as is shown in Fig. 18.3. This smoothed curve, familiar to mathematicians and biologists is also called the Gaassion curve after the great German mathematician Gaass, who investigated its properties and wrote the equation for it.

It should be noted that the term normal as here used is simply a name for this particular curve and does not have any of the usual connotations as, for example, in speaking of "a normal child". Normal as a technical term in statistics does not mean free from abnormalities. Similarly the term *normal distribution* means any frequency distribution whose form corresponds to that of the normal curve.

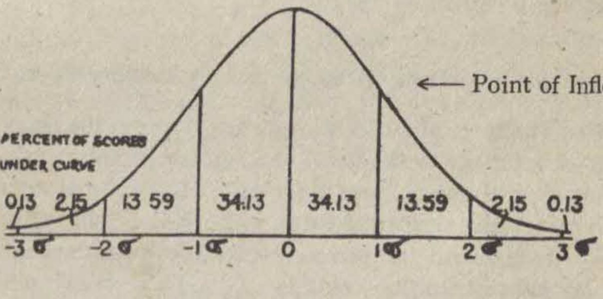


Fig. 18.3 : Normal Probability Curve.

The principle of probability is used in many situations. In lotteries, prizes are distributed through the system of lucky numbers. In races we find the same principle. A particular horse cannot always win and be a lucky number on all occasions. Life insurance companies take into consideration the probability of an individual's life span at various age limits and then work out the amount of the premium. We know that the probability of our living for the next twenty years generally decreases with our age, hence the amount of premium for the age groups say, 20, 25, 30, 40, 45 will gradually go on increasing. In all the above illustrations, the probability of an event that is, the expected frequency of occurrence of the event is taken into consideration.

Many human characteristics such as height, weight, strength of grip, learning ability, mechanical ability, co-operativeness, social dominance etc., yield normally distributed measurements like these. Some traits may not be distributed in this form; thus the human skin or eye colour is not distributed normally because people fall into fairly distinct groups on the basis of these traits. Human height is a more continuous trait and it may be distributed in a form similar to the normal curve. All the above distributions give *a bell-shaped curve* which is known as *the normal curve*. The normality of data is a concept of great usefulness in statistical theory and practice and no student can use and interpret statistics successfully without some understanding of the normal curve. It is no exaggeration to say that the *normal law* namely, the greater a deviation from the mean, the less frequently it occurs, is the foundation of statistical theory.

What are the properties of the Normal Probability Curve ?

(1) The normal curve is symmetrical about the mean. The number of cases below the mean in a normal distribution is equal to the number of cases above the mean. Hence the mean and the median coincide in a normal curve. (2) The height of the curve is at its maximum at the mean. Hence the mean and the mode of the normal distribution coincide. (3) There is one maximum point of the normal curve, which occurs at the mean. The height of the curve declines as we go in either direction from the mean. This dropping off is slow at first, then rapid and then slow again. Theoretically the curve never touches the base line. Its tails approach but never reach the base line. Hence the range is unlimited. (4) The points of inflection i.e., the points where the curvature changes in direction are each ± 1 SD from the mean ordinate, (Fig. 18.3). (5) The height of the curve at a distance of 1σ from the mean is 60·7 percent of the height at the mean. The height of the curve is shown by drawing the ordinates at given distances. The height of the curve at 2σ and 3σ distances from the mean is respectively 13·5 percent and 1·1 percent of the height at the mean.

Area Relationships under the Normal Curve

Perhaps the greatest usefulness of the normal curve lies in the relationship of the amount of area under the curve lying between certain limits on the base line. In terms of test scores, for example, this simply means the number or percentage of the cases to be expected between two score points. Since the ordinate at a given sigma distance from the mean of the normal curve always has the same relationship to the mean ordinate, it follows that the area under the curve included between the mean ordinate and an ordinate a given sigma distance from the mean is always the same proportion of the total area under the curve. If we erect perpendiculars from the base line at the mean and at a point 1 SD from the mean, the area included between these perpendiculars will always be 34·13 percent of the total area. To say the same thing in another way, from the mean to + 1σ are to be expected about 34·13 percent of the cases and from the mean to – 1σ, another 34·13 percent of the cases (Fig. 18.3). If there were a normal distribution with 10,000 cases, 3413 of them would be expected between the mean and 1σ. In terms of percentages, it would be 34·13 percent or 34·13 cases in 100. The total interval from + 1σ to – 1σ contains 68·26 percent of the cases. Similarly 47·72 percent of the total area will be included between the mean ordinate and an ordinate 2σ from the mean and 49·87 percent of the total area will be subtended between the mean and a point 3σ from the mean. Similar statements for ordinates at other distances from the mean apply alike to any normal curve.

Applications of the Normal Probability Curve

The above relationship may be employed to derive a number of important types of information about any normal distribution of measures. Following is an illustration.

To determine the percentage of cases in a normal distribution within given limits.

Illustration : Given a normal distribution of intelligence test scores, with a mean = 100, and S.D. = 16. (a) What percentage of the cases fall between an IQ of 84 and 116 ? (b) What percentage of the cases lie above an IQ of 132 ? (c) Below an IQ of 92 ?

(a) An IQ of 116 is 16 points above the mean and an IQ of 84 is 16 points below the mean. If we divide this scale distance of 16 points by the σ of the distribution (i.e. by 16), it is clear that 116 is 1σ above the mean and that 84 is 1σ below the mean. There are 68·26% of the cases in a normal distribution between the mean and $\pm 1\sigma$. Hence 68·26% of the scores in our distribution fall between IQ's of 84 and 116 (Fig. 18.4).

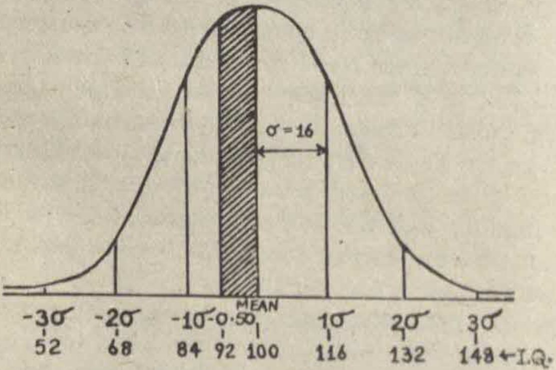


Fig. 18.4 : Normal Probability Curve.

(b) An IQ of 132 is 32 score units above the mean or 32/16 = 2σ units above the mean. It was shown that, 47·72% of the cases lie between the mean and this point, that is between 100 and 132. Since the distribution is symmetrical, 50 percent of the cases will lie above the mean. Hence 50 – 47·72 = 2·28 percent of the group will have their IQ above 132. If Jayant has an IQ of 132, only 2·28 percent of the students from the whole group have their IQ's greater than that of Jayant.

(c) An IQ of 92 is below the mean hence we would first find the percent between this score and the mean, and then subtract this percentage from 50%, the total percent below the mean. The score of 92 is $-0{\cdot}5\sigma$ from the mean. Between the mean and $-0{\cdot}5\sigma$ are 19·15 percent of the cases (Fig. 18.4). Hence 50 − 19·15 = 30·85 percent of the cases lie below an IQ of 92. If Geeta has an IQ of 92 she surpasses 30·85 percent of the students from the whole group.

Significance of the Normal Curve

Distributions of many physical traits exhibit the bell-shaped form; hence it seems probable that the same would also be true for many mental traits. It is dangerous, however, to argue this by analogy from one type of trait to another. We are not justified in talking loosely about any underlying 'law of normality' as if such a law applied to the distribution of measures of any trait regardless of the character of the population considered. Perhaps one of the principal reasons that we have exaggerated the importance of the normal curve in education and psychology is the fact that scores obtained on educational and psychological tests for almost any unselected group of pupils, so frequently present a bell-shaped form of distribution. This, however, is not of any very fundamental significance, since most of these tests are deliberately constructed so as to yield approximately symmetrical distributions of scores. In nearly all test construction the author tries to adjust the difficulty of the items in such a way so as to get a normal distribution of scores. If he desired to do so, the test author could as easily prepare a test that would yield a distribution markedly skewed.

In dealing with the complex variables of social sciences, it is rarely if ever possible to demonstrate *a priori* that the conditions of normal distribution are present. For example, consider the trait *musical ability* or *mechanical skill*. Do these traits exhibit a normal distribution ? We cannot prove it because the factors which determine these abilities are too little known to justify the assumption that they combine in the same proportions as do the head and tail combinations in 'chance' distributions of coins.

Hence normality or non-normality must be inferred from the distribution of test scores themselves. There is nothing wrong in making such an inference, provided we recognize the limitations involved. The resulting symmetrical distribution of test scores should not be taken as a conclusive proof of *normality* of the trait being measured. We prefer the normal curve in test construction because this type of distribution generally fits the data better and such a curve because of its properties, is more useful in interpreting the test results. But it should be remembered that the theoretical justification and empirical use of the normal curve are two quite different matters.

Skewness

What do you understand by the term skewness? What are the types of skewness? How does skewness differ from dispersion?

In a frequency polygon or a histogram of test scores, usually the first thing which strikes the eye is the symmetry or lack of it in the figure. *A distribution is said to be skewed when the mean and the median fall at different points in the distribution and the balance or centre of gravity is shifted to one side or the other.* The word skewed means 'lacking symmetry' or 'distorted'. Skewness depends upon the manner in which the scores in a series scatter about the average value. When the scatter is greater on one side of the point of central tendency than on the other, the distribution is skewed.

In a normal distribution, the mean equals the median exactly and the skewness is, of course, zero. The more nearly the distribution approaches the normal form, the closer together are the mean and the median, and the less the skewness. Distributions are said to be skewed *negatively* or to the *left* when scores are massed at the high end of the scale (the right end) and are spread out more gradually towards the low end (or left) as shown in Fig. 18.6. Distributions are skewed *positively* when there is a piling up of scores at the low end and a long tail running up into the high scores, Fig. 18.5. The median lies in the middle while the mean and the mode remain on opposite sides of the median. Note that the mean is pulled more toward the skewed end of the distribution than is the median. In fact greater the gap between the mean and the median, greater the skewness. Moreover, when skewness is negative, the mean lies to the left of the median and when skewness is positive the mean lies to the right of median.

Dispersion gives us an idea of the deviation of scores from the mean or from the median. The measures of dispersion namely σ and Q fail to show the *manner of distribution* of the deviation. Skewness on the other hand, shows the position of the mean in relation to the mode or the median and brings out the distortion of the curve. It shows the direction of asymmetry.

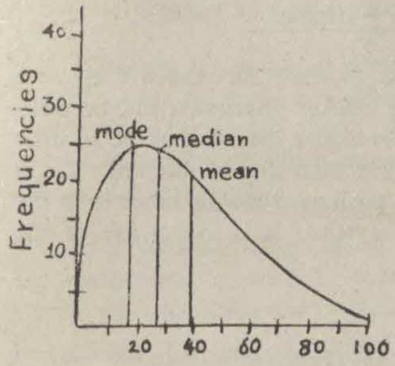


Fig. 18.5 : **Positive Skewness.**

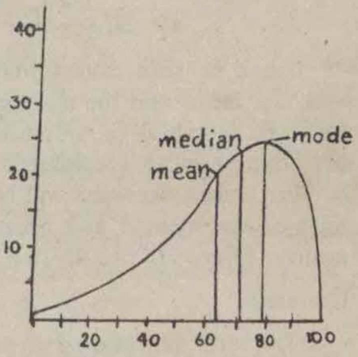


Fig. 18.6 : Negative Skewness.

What are the different tests for detecting skewness? How will you measure it?

The presence of skewness can be detected in several ways. The lack of symmetry in a series is shown :

(i) When the mean, the median and the mode do not coincide, or

(ii) When the sum of the positive deviations from the median in a series is not equal to the sum of the negative deviations, or

(iii) When the quartile points Q_1 and Q_3 are not equidistant from the median, or

(iv) When at points of equal deviations on either side of the mode, the frequencies are unequal, or

(v) When the frequency curve drawn is not a bell-shaped curve but rather its base is stretched to a greater extent on either side (Figs. 18.5 and 18.6).

The second and the third tests mentioned above do not provide a conclusive proof as to the skewness of the distribution hence the first and the fourth tests are generally used. The fifth one is some what laborious. However to know how much does the curve lack symmetry, we must have some *measure of skewness.*

Measures of Skewness : The methods mentioned above are all crude and they fail to give a numerical measure of skewness. They are useful only in preliminary analysis. One of the widely used measures of skewness is given by the formula :

$$Sk = \frac{3\,(\text{mean} - \text{median})}{\sigma} \quad \ldots \quad (19)$$

It will be seen from formula 19, that skewness will be zero when the mean and the median coincide. Skewness will be positive when the value of the mean is higher than that of the median and negative when the value is less than that of median. Stated in other words, skewness will be positive when the mean is to the right of the median and negative when it is to the left of the median (Figs. 18.5 and 18.6).

Kurtosis

The term *kurtosis* refers to the 'peakedness' or 'flatness' of a frequency distribution as compared with the normal. When

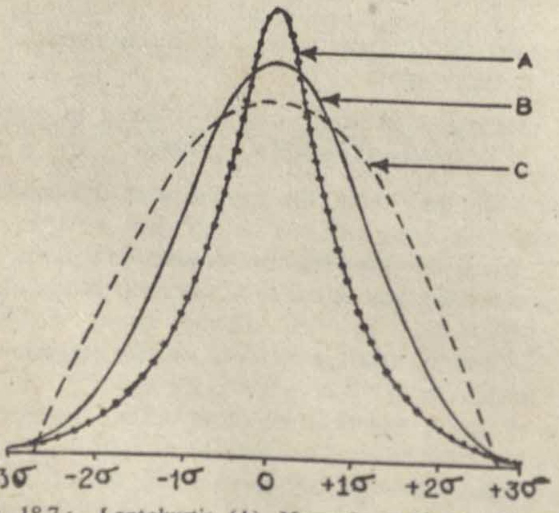


Fig. 18.7 : Leptokurtic (A), Normal or Mesokurtic (B) and Platykurtic (C) Curves.

there is a high concentration of scores in the neighbourhood of the point of central tendency, the distribution is relatively narrow across the shoulders. Relatively high and narrow distributions are described as *leptokurtic*. When there is low concentration of scores in the neighbourhood of the central tendency, the distribution is relatively broad across the shoulders. Such relatively flat topped distributions are described as *platykurtic*. A normal distribution is called *mesokurtic*, (Mesos means middle or medium). Fig. 18.7 shows these three types plotted on the same diagram around the same mean.

What are the causes of skewness and kurtosis ? : It is often important for the research worker to know why his distributions diverge from the normal form. Following are the common causes of asymmetry, all of which must be taken into consideration by the careful experimenter.

1. *Selection* : Selection is a potent cause of skewness. If the sample you choose is a biased one, the distribution of the scores will not exhibit the bell-shaped form.

2. *Unsuitable or Poorly Made Tests* : Normality or lack of normality is dependent upon the number of items and their difficulty. If a test is too easy, scores will pile up at the high score end of the scale and will give negative skewness whereas if the test is too hard, scores will pile up at the low score end of the scale giving a positively skewed curve.

3. *Non-Normal Distributions* : Skewness or kurtosis or both will appear when there is a real lack of normality in the trait being measured, e.g., if a loaded dice is tossed for a number of times, the resulting distribution will certainly be skewed and probably be peaked. This is because the loaded side is a dominant factor in determining the result of the tosses. Non-normal curves often occur in medical statistics. In the case of a childhood disease, for example, death rate would be maximum in the early ages and the rate would decrease with increase in age. The distribution would be positively skewed.

4. *Errors in the use of Tests* : Errors in timing or in giving instructions, errors in scoring, differences in motivation, all of

these factors, if they cause some students to score higher and others to score lower than they normally would, tend to make for skewness in the distribution.

Other Forms of Distribution

A Bimodal Curve : Such a curve is obtained when there are two distinct peaks in a distribution. This is possible in a class which has two distinct groups of students such as bright students and dull ones, (Chapter 15, Table 7) If you take a distribution of heights of a group of college students which includes half Japanese and half Americans, the resulting curve might have two peaks, that is, it will be bi-modal, with one peak at the average of American height and the other at the average of Japanese height, as in Fig. 18.8. Sometimes the distribution may have more than two peaks; in that case we get a multimodal curve.

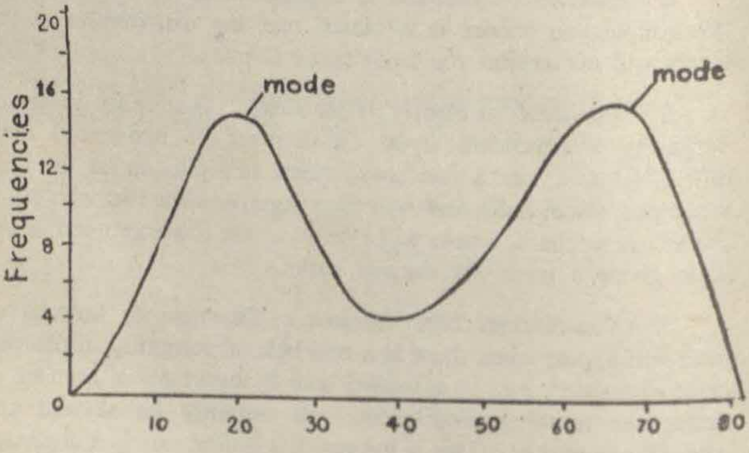


Fig. 18.8 : Bimodal Curve.

J-Type Curve : If a curve be drawn showing the income of Indian people, we get a distribution which exhibits a J-type curve (Fig. 18.9). There are many people with small and moderate income while only a few with very large incomes. Such a distribution of incomes will be positively skewed. A J-shaped curve is essentially non-normal.

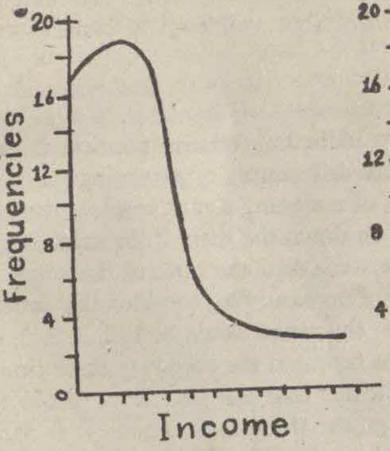


Fig. 18.9 : J-Type curve.

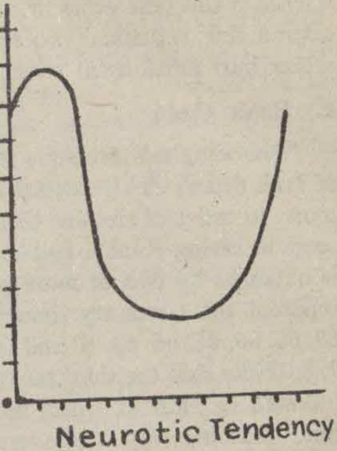


Fig. 18.10 : U-Type curve.

U-Type Curve : Another type of non-normal distribution is the U-shaped curve shown in Fig. 18.10. U-shaped curves are rarely encountered in mental and physical measurements. Such a distribution is possible when there are two distinct groups in a class, e.g., if a *Neurotic Inventory* is given to a group composed about equally of normals and of mentally ill persons, the normals will tend to make low scores while the abnormals will tend to make high scores. This makes for a dip in the centre of the distribution. Similarly, if we plot the distribution showing the probability of death at various are levels, we shall get a U-Type curve.

Making Test Scores Comparable

When an answerbook has been valued as per instructions, the score obtained is called a *raw* score or a *crude* score. These raw scores tell us nothing more than the number of answers marked correct on each test. A raw score by itself means very little. A raw score of 40 on one test is not directly comparable to a raw score of 45 or 50 on another test. The difficulty is that the units are not comparable. Mary has a score of 56 in science. What does this score signify ? She has a score of 42 in arithmetic and a score of 34 in reading. How can you compare her standing in the three subjects ? A novice may say that her performance in science is the best while that in reading is the worst. Let us consider a few statistical procedures, that will help to convert raw scores into meaningful terms.

1. Rank Order

The simplest technique for indicating relative position is that of rank order. This method consists simply of arranging the new scores in order of size and then of assigning a rank to each, the top score receiving Rank 1 and so on down the line. If the same score is obtained by two or more persons then the rank of the score is repeated the necessary times. For example, consider the scores 69, 68, 66, 66, 66, 62, 60 and 56; the ranks would be 1, 2, 3, 3, 3, 6, 7, 8. Note that the third rank is repeated the necessary three times followed by rank 6, which indicates that the next score is 6th in line. Another way is to average the three ranks $(3 + 4 + 5)/3$ and give each individual a rank of 4. The next rank will be 6. Obviously if the number of scores is large the process of ranking will become unwieldy. Moreover to know just the rank of a score is not sufficient because the same rank would have different meaning in groups of different sizes, for example a rank of 15 in a group of 40 is not the same as a rank of 15 in a group of 120. For this reason ranks are ordinarily expressed in relative terms as *percentile ranks*.

2. Standard Scores

Another method of indicating a pupil's relative position in a group is by showing how far his raw score is above or below the mean. This is the approach used with *standard scores*. Basically standard scores express test performance in terms of standard deviation units from the mean. Such deviation scores are called *sigma scores*. There are numerous types of standard scores used in testing. They are all based on the same principle.

(a) **Z-scores:** A *z*-score is defined as the deviation of a score from the arithmetic mean in *standard deviation units*. Ordinarily the letter *z* is used to refer to this type of *standard score*. The formula for the *z*-score is:

$$\text{Standard score } (z) = \frac{X - M}{\sigma} = \frac{x}{\sigma} \qquad \dots \quad (20)$$

where X = any raw score; M = mean of the raws scores

x = deviation from the mean; σ = standard deviation.

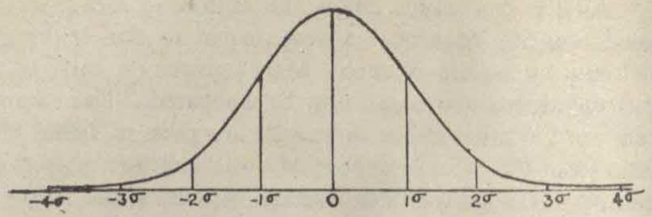
All the raw scores below the arithmetic mean will thus become negative *z*-scores. A score equal to the arithmetic mean will equal a *z*-score of zero. Mary's scores on each of the three tests mentioned above can now be compared. The mean and the standard deviation of the three tests are given in Table 18.1. It is clear from the tabulation that Mary is well above average in the reading test, slightly above average in science and below average in arithmetic.

Table 18.1 : Converting raw scores into standard scores

	Raw Score	Mean M	$x = X - M$	σ	*z* score	T-Score
Science	56	52	4	8	0·5	55
Arithmetic	42	47	–5	5	– 1·0	40
Reading	34	28	6	4	1·5	65

To illustrate further the value of standard deviation units for expressing relative position in a group, raw scores of the above three tests have been placed beneath the row of deviations along the baseline of the curve in Fig. 18.11. Note that the mean raw scores of these tests have been placed at the zero point on the base line of the curve. Note that a *z*-score of *one* is equivalent to a score of 60 (52 + 8) in science, a score of 52 (47 + 5) in arithmetic and a score of 32 (28 + 4) in reading. If we convert all of the raw scores on the three tests to the standard deviation units in this manner, it is possible to compare directly the performance on these tests. For example, a raw scores of 68 in science and 57 in arithmetic are equal because both are 2 units above the mean. When this conversion to *z*-score is made, the raw scores are no longer needed.

This method permits us to convert raw scores to a common scale which has equal units and which can be readily interpreted in terms of the normal curve. It is apparent that σ scores are simply plus or minus deviations from the mean expressed in σ



zScore	-4	-3	-2	-1	0	+1	+2	+3	+4
Science	20	28	36	44	52	60	68	76	84
Arithmetic	27	32	37	42	47	52	57	62	67
Reading	12	16	20	24	28	32	36	40	44

Fig. 18.11 : Normal curve and the standard scores.

units. With the help of standard scores we can compare the standing of two or more students in various tests. Consider the scores of Judy and Mary :

	Science	Arithmetic	Reading	Total	Remarks
Mary	56	42	34	132	Mary's total score is higher than that of Judy
Judy	48	42	40	130	

Convert these raw scores into *z*-scores by using formula (20) :

	Science	Arithmetic	Reading	Total of *z*-scores	
Mary	+0·5	−1·0	+1·5	1·0	Judy's score is higher than that of Mary
Judy	−0·5	−1·0	+3·0	1·5	

When put into standard score form, Mary's and Judy's scores can be compared directly. When the raw scores are compared, Mary's score appears to be higher but with standard scores, Judy's performance appears to be better. The standard scores give a correct picture because they have a common scale which has equal units.

A practical disadvantage of such scores is the fact that they require the use of decimal points to express the z-scores. For these reasons z-scores are usually converted into a new distribution with M and σ so selected as to make all scores positive and relatively easy to handle. Such scores are also called *standard* scores. One of the types of standard scores is the T-score.

(b) *T-Scores* : T-scores serve the same purpose as z-scores and are based on the same principle. They have the advantage of being always positive and expressed in larger units thus removing the necessity of dealing with negative numbers and decimal fractions. T-scores are obtained by multiplying the z-scores by 10 and adding the product to 50. In other words the mean is set at 50 and the S.D. is taken as 10 in converting z-scores into T-scores. Thus

$$\text{T-score} = 50 + 10\,(z) \quad \ldots \ldots \quad (21)$$

A T-scoore of 60 means a score of one standard deviation above the mean. A z-score of 1·4 would be a T-score of 64. In Table 18.1, we find Mary's z-score of 4·5 equivalent to a T-score of 55 and a z-score of – 1·0 is equivalent to a T-score of 40, while Judy's z-scores of – 0·5 and 3·0 are equivalent to T-scores of 45 and 80 (Table 18.1).

(c) *Deviation IQ* : Another widely used standard score is the deviation IQ. Here the mean is set at 100 and the standard deviation at 16. Thus a pupil with an IQ of 84 has scored one standard deviation below the mean and a pupil with an IQ of 116 has scored one standard deviation above the mean (Fig. 18.12).

3. Percentile Ranks

A third method of making test scores comparable is by computing percentile ranks. A *percentile rank* indicates a pupil's relative position in a group in terms of the percentage of pupils scoring below him. The points that divide a distribution into 100 equal divisions or percents are called *percentiles* or more simply *centiles*. The median is the 50th percentile, since 50 percent of all the measures lie below this point. Similarly Q_1 is the 25th percentile and Q_3, the 75th percentile. If Bob has a score higher than that made by 60 percent of the students in a group, we would say that he is at the 60th percentile. In general then the *p*th percentile in a distribution of scores may be defined as that point on the scale below which *p* percent of the cases fall.

The boundaries of the percentile scale are indicated by the two percentiles 100 and 0. The zero percentile denotes the point below which lie zero percent of cases. P_0 marks the exact lower limit of the lowest class interval while P_{100} marks the exact upper limit of the highest interval. These two percentiles represent limiting points. Their principal value is to indicate the boundaries of the percentile scale. Note that P_0 does not correspond to a score of zero and P_{100} does not correspond to a perfect score. Percentiles should not be confused with the common percentage scores. Lata obtained 70 percent of marks; here the raw score is expressed in terms of the percentage of correct items, while percentiles are derived scores in terms of percentage of individuals.

(a) *Computation of Percentiles* : We find percentile equivalents for each raw score by arranging the raw scores in order of decreasing magnitude and finding for each, the number of scores lying below it. Each of these scores we then divide by the total number of scores and multiply the quotient by 100. Suppose for example, out of 44 cases 11 cases fall below the raw score of 66; it means 25 percent of the scores fall below 66. Then the percentile rank (PR) of the raw score 66 for this particular group is 25 which is designated as P_{25}.

$$P_{25} = 66$$

It means that 25th percentile is 66, that is, there are 25 percent of cases scoring below 66. Or we can say that Johnny who scored 66, has a percentile rank of 25. For computing the percentiles of scores given in the form of a frequency table, we use the same method as was employed in computing the median. The formula is :

$$P_p = l + \left(\frac{{}_pN - F}{f}\right) \times i$$

P_p = percentile required, such as P_{10}, P_{16} or P_{50}
l = exact lower limit of the class interval containing P_p
${}_pN$ = part of N to be counted off in order to reach P_p
F = sum of all the scores below the class interval containing P_p
f = number of scores in the class containing P_p
i = class interval

In Table 18.2, the percentile points P_5, P_{10}, P_{90} have been computed by the formula given above. We may illustrate the method with P_{10}, tenth percentile. We want to know the point below which lie 10 percent of the scores of the given group. Here ${}_pN$ i.e. 10 percent of $N = 10/100 \times 80 = 8$. From the cumulative frequency as shown in Table 18.2, we find that P_{10} falls in the interval 70–79. Substituting the values;

${}_pN = 8$, $F = 3$, $l = 69{\cdot}5$, $f = 7$ and $i = 10$

we find that $P_{10} = 76{\cdot}64$.

It means that 10% of the students scored less than 76·64. The other percentile values are found in exactly the same way as P_{10}.

Table 18.2 : Computation of Percentiles

1 Scores	2 f	3 Cumulative frequency	4 Determine P_0, P_5, P_{10}, P_{45}, P_{50}, P_{90}, P_{100}
130—139	2	80	5% of 80 = 4 10% of 80 = 8 45% of 80 = 36
120—129	7	78	50% of 80 = 40 90% of 80 = 72
110—119	11	71	$P_0 = 59\cdot5$
100—109	14	60	$P_5 = 69\cdot5 + \left(\frac{4-3}{7}\right) \times 10 = 70\cdot9$
90—99	20	46	$P_{10} = 69\cdot5 + \frac{(8-3)}{7} \times 10 = 76\cdot64$
80—89	16	26	$P_{45} = 89\cdot5 + \frac{(36-26)}{20} \times 10 = 94\cdot5$
70—79	7	10	$P_{50} = 89\cdot5 + \frac{(40-26)}{20} \times 10 = 96\cdot5$
60—69	3	3	$P_{90} = 119\cdot5 + \frac{(72-71)}{7} \times 10$ $= 120\cdot9$
	N = 80		$P_{100} = 139\cdot5$

(b) *Computation of Percentile Ranks* : In Table 18.2 we find that 45th percentile (P_{45}) is 94·5 which means that there are 45 percent of cases scoring below 94·5. On the other hand we can say that Barbara who has a score of 94·5 has a percentile rank of 45. The *percentile rank* (PR) of a given score is the number representing the percent of cases in the total group, lying below the given score while the *percentile* is the score below which lie a given percent of cases. The distinction between percentile and percentile rank will be clear if the reader remembers that in calculating percentiles he starts with a certain percent of N, say 10% or 70%. He then counts into the distribution the given percent and the point reached is the required percentile, P_{10} or P_{70}. The procedure for computing percentile ranks is the reverse of this process. Here we begin with an individual score and determine the percentage of scores which lie below it. If this percentage is say 20, the score has a percentile rank or PR of 20 on a scale of 100.

Let us determine the percentile rank (PR) of a score of 92 in the distribution of scores in Table 18.2. This score falls in the class interval 90–99, hence it clearly exceeds the 26 cases falling below that interval. To find how many cases lie below 92 in the interval 90–99 we assume that the 20 scores in that class are evenly distributed over the interval. The score 92 is 2·5 score units from 89·5, the exact lower limit of the interval 90–99, see Fig. 18.12.

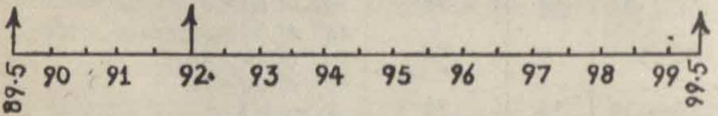


Fig. 18.12: Enlargement of the interval 90–99

The number of scores that lie below 92, in the interval 90–99 is $\frac{2\cdot5\times20}{10} = 5$. Hence the total number of scores below 92 will be 26 + 5 = 31. There are 31 scores below 92, in this group of 80, (N = 80). Therefore the percentile rank of score 92 is

$$\text{PR of } 92 = \frac{31}{80} \times 100 = \frac{155}{4} = 38\cdot75.$$

The percentile ranks for several scores can be read directly from Table 18.2. For example, there are 60 scores that lie below 109·5. The PR of 109·5 is $\frac{60}{80} \times 100 = 75\cdot0$, while the PR of 84·5 is 22·5.

$P_{45} = 94{\cdot}5 \quad \therefore$ PR of $94{\cdot}5 = 45$
$P_0 = 59{\cdot}5 \quad \therefore$ PR of $59{\cdot}5 = 0$
$P_{90} = 120{\cdot}9 \quad \therefore$ PR of $120{\cdot}9 = 90$.

(c) *Uses of Percentiles* : The most frequent application of percentiles is in testing. Standardized tests usually report norms in terms of the percentile values of raw scores. The percentile system enables us to compare the standing of a student on various tests. Knowing that Karen has scored 52 on an arithmetic test and 55 on a science test does not help us in estimating her ability. However if we know that the two scores correspond to 84th and 69th percentiles we can compare her standing in the two subjects. Similarly percentiles are useful in comparing one individual with the rest of the group. Percentile norms are suitable for any type of test such as an aptitude or an achievement test or a personality inventory.

(d) *Limitations of Percentiles* : (1) Percentile norms will have no meaning if the norm group is not taken into consideration. Marvin has a percentile rank of 60 on a vocabulary test. This does not convey much meaning unless we state the reference group. If Marvin is in the VIII class and if the above PR is with reference to students of IX standard, his performance may be considered as superior, but if the reference group is of VI standard, Marvin's performance may be considered as below average. (2) The percentile rank units are of unequal length. Note the inequality in Fig. 18.13. The difference between the 10th and 20th percentiles is not the same as that between the 50th and 60th percentiles. This is mainly because the test scores usually pile up in the middle and tail off at either end. (3) Because of the inequality of units, it is not correct to combine the percentile ranks of a child on two or more tests to find the average percentile rank. (4) Percentiles are not well suited to the computation of means, correlations and other statistical measures.

Comparison of the different Score Systems

The student studying the various score systems will find that many of the methods of deriving meaningful scores are anchored to the dimensions and characteristics of the normal curve (Fig. 18.13).

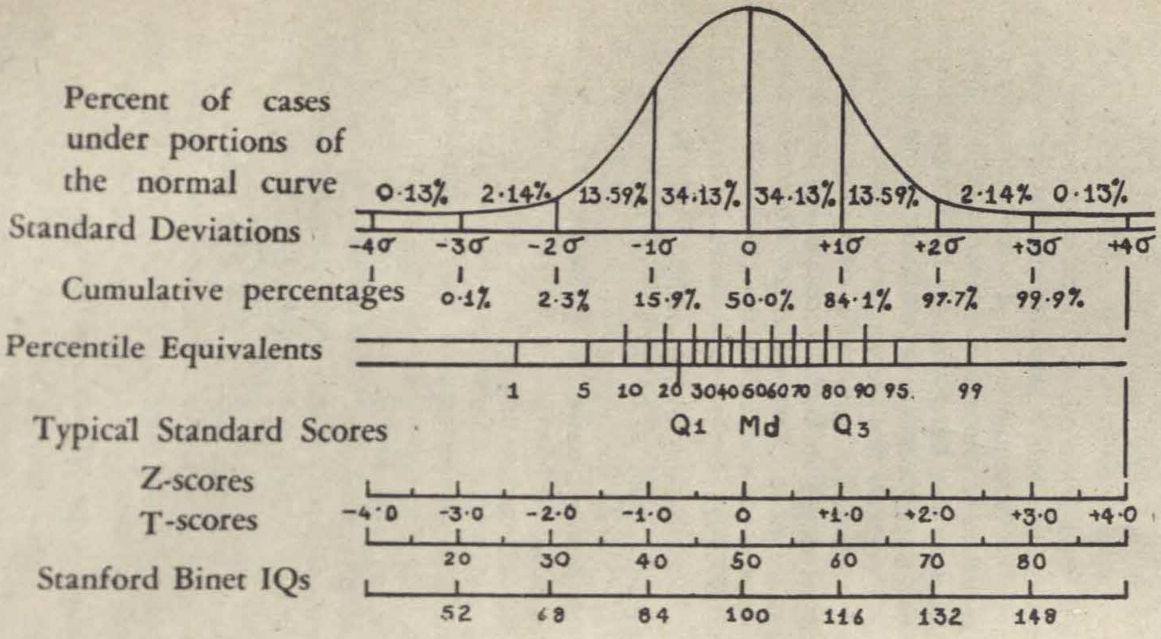


Fig. 18.13 : Comparison of different score systems

Fig. 18.13 illustrates the interrelated nature of the various scales for reporting relative positions in a normally distributed group. A raw score of one S.D. below the mean, for example, can be expressed as a Z-score of – 1·0, a percentile of 16, a T-score or 40, or a deviation IQ of 84. Thus the various score systems are merely different ways of saying the same thing and we can readily convert back and forth from one scale to another, provided of course that we get a normal distribution of scores.

FOR FURTHER STUDY

1. Suppose a 50 item test of low difficulty was administered to a group of 100 students. What type of distribution of scores is likely to occur? How would the distribution be, if the items were of high difficulty?
2. A fifty-item vocabulary test given to 150 pupils, yielded scores ranging from 18 to 50. Ninety-seven pupils scored between 40 and 50. What would this distribution of scores look like? What could you say about the suitability of the test for the group? What measure of central tendency would be most suitable? Why? What measure of variability would you probably use?
3. The mean and standard deviation of a normal distribution of 500 scores are 75·0 and 15·0 respectively.
   (a) How many scores lie between 45 and 90?
   (b) How many scores lie below 60·0?
4. Show the positions of the mean, the median and the mode in a positively and a negatively skewed curve.
5. Explain the causes of skewness in the distribution of scores.
6. (a) Explain the properties of a normal probability curve.
   (b) Mention five characteristics or traits that give a normal distribution.
   (c) Explain the uses of a normal curve in interpreting test scores.
7. Compute the percentiles P_{20}, P_{25}, P_{75} and P_{80} for the distribution given in Table 18.2 of this chapter. Interpret the results.
8. Compare the achievement of Brenda on the following tests :

	English	Georgraphy
Brenda's raw score	39	65
Mean score of the class	30	50
S.D.	6	15

(a) Compare her z-scores; in which test is her performance better ?
(b) Convert her raw scores into T-scores.
(c) If Gloria has a z-score of $-0{\cdot}5$ in English and a z-score of $-0{\cdot}6$ in Georgraphy, what is her raw score in each subject ?
(d) Compare the performance of Brenda with that of Gloria.

9. Toss six coins 64 times and record the number of heads and tails after each throw. Plot frequency polygons of obtained and expected occurrences on the same axes. Compare the means and S.D.'s of obtained and expected distributions.

10. *Comment on the following statements* :

    (i) Scores on a mechanical aptitude test present a bell shaped curve, hence the trait is normally distributed.
    (ii) If a test is too difficult for a certain large group, the distribution of scores will present a skewed curve.
    (iii) If two distributions have the same means and medians, they have the same skewness.
    (iv) Greater the gap between the mean and the median, greater is the skewness.
    (v) A percentile of 0 corresponds to a score of zero.
    (vi) If the median is greater than the mean, the distribution of scores is skewed to the left.
    (vii) If two distributions have equal means and equal medians, the more variable of the two distributions has the greater skewness.
    (viii) Authors of various tests prefer the normal curve in test construction.

11. *Explain the meaning of the following statements* :

    (i) The fortieth percentile (P_{40}) on a certain test is 37.
    (ii) Percentile is not the same as the common percentage score.
    (iii) z-scores and T-scores are basically the same.
    (iv) Percentile rank units are of unequal length.
    (v) Standard scores help us to compare the standing of a student on different tests.
    (vi) Tanuja has a percentile rank of 71 on a test in arithmetic.

12. Compute measures of skewness for the two distributions in chapter 16, problem 11, page 329.

13. In a sample of 1000 cases, the mean score on a certain achievement test is 100 and S.D. is 20; assuming normality of distribution.

    (i) What percent of scores lie between 90 and 120 ?
    (ii) How many individuals score below 80 ?
    (iii) How many students score between 80 and 120 ?
    (iv) If Aditya has a score of 120 on this test, what is his percentile rank ?

14. Margaret achieves a score of 56 on a reading test and a score of 45 on a geography test. How will you compare her standing on the two tests with the help of the following information? In which test is her performance better, in regard to the rest of the class?

	Reading	Geography
Mean for the whole class	56	40
S.D. for the whole class	8	5

15. An intelligence test is given to a large group of pre-university students. The mean raw score is 60 and S.D. is 12. What inference can you draw about the standing of the following students if the distribution of scores is normal?

Jayant	72	Meena	60	Karen	36
Lata	66	Jim	48	Mike	84

16. Convert Jayant's and Lata's raw score into (i) *z* and T scores, (ii) deviation IQ's and (iii) percentiles.

17. Illustrate with the help of a normal curve, how various types of scores can be compared with one another.

Answers

1. Negatively skewed distribution when the items are easy.
2. (i) a negatively skewed distribution (ii) the test appears to be easy for the group (iii) median and Q would be suitable measures.
3. (a) 409 (b) 79 (round numbers)
7. $P_{20} = 83{\cdot}25$ $P_{25} = 85{\cdot}75$ $P_{75} = 109{\cdot}5$ $P_{80} = 113{\cdot}14$

8.	English	Geography	
Brenda's *z*-scores	1·5	1·00	Brenda's performance is better than that of Gloria.
" T-scores	65	60	
Gloria's raw scores	27	41	
" *z*-scores	–0·5	–0·6	

10. (i) cannot definitely infer (ii) a positively skewed curve (iii) not necessarily true (iv) true (v) false (vi) true (vii) false (viii) It is convenient to interpret the results.
13. (i) 53·28 (ii) 159 (iii) 683 (round number) (iv) PR 84 13.
14. Margaret's performance in geography is better than that in reading. Compare the *z* scores on the two tests.
15. Mike, Jayant and Lata have IQ higher than average. Meena has an average IQ, while Jim and Karen have IQ below average.

16.	Raw score	*z*	T	deviation IQ	Percentile
Jayant	72	+ 1·0	60	116	84
Lata	66	+ 0·5	55	108	69

* * *

APPENDIX I

Objective based Test Items

We have discussed in chapter 11 that test items should be objective-based. Here an attempt is made to show how items can be prepared to assess various objectives. While writing down items, learning outcomes should also be indicated.

Subject : Mathematics

Directions : Each statement of a problem is followed by four responses A, B, C and D. Put within the bracket at the end of each problem, the letter indicating the correct answer.

1. *Objective* : Understanding *Topic* : Ratio and Proportion

Learning Outcome : Finds relationship.

Q: If $\frac{2}{3}$ of a number is 10, then the number is ———

A. $3\frac{1}{3}$ B. 18 C. $6\frac{2}{3}$ D. 15 ... ... ( )

2. *Objective* : Understanding *Topic* : Problems on equations.

Learning Outcome : Translates verbal statements into symbolic ones.

Q : A number of two digits is $10x + y$. The number obtained by interchanging the digits is

A. $y+40x$ B. $10y+x$ C. $\frac{1}{x}+\frac{1}{y}$ D. $\frac{10}{x}+\frac{1}{y}$ ... ( )

3. *Objective* : Understanding *Topic* : Fractions

Learning Outcome : Discriminates.

Q : Which of the following statements is false ?

A. $\frac{2}{10}=\frac{7}{35}$ B. $\frac{2}{10}=\frac{2\times 7}{10\times 35}$

C. $\frac{2}{7}=\frac{2+10}{7+35}$ D. $\frac{2+7}{7}=\frac{10+35}{35}$ ... ... ( )

4. *Objective* : Understanding *Topic* : Fundamental operations

Q : Which number must be placed in the box to make the following statement true? $6 \times 2 = (\square \times 2) + (1 \times 2)$

A. 7 B. 6 C. 5 D. 4 .. ( )

5. *Objective* : Application *Topic* : Percentage

Learning Outcome : Selects appropriate method.

Q : The population of a town increases by 10% every year over the population of the previous year. If the population in 1970 is 17,000, what would be the population in 1971 ?

A. $17{,}000 \times \frac{100}{110}$ B. $17{,}000 \times \frac{110}{100}$

C. $17{,}000 \times \frac{90}{100}$ D. $17{,}000 \times \frac{100}{90}$ ... ( )

6. *Objective* : Knowledge *Topic* : Laws of operation

Learning Outcome : Recalls laws of operation.

Q : Which of the following statements is not true ?

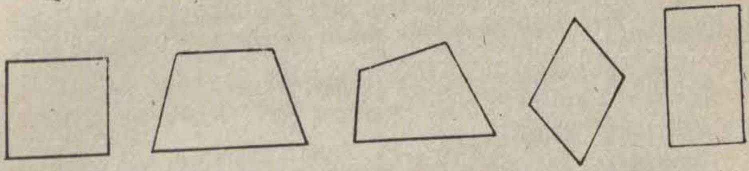
A. $0 \times 176 = 421 \times 0$ B. $52 \times 21 = 21 \times 52$

C. $44 \div 44 = 9 \div 9$ D. $13 \times 32 = 31 \times 23$ .. ( )

7. *Objective* : Knowledge *Topic* : Geometrical Figures

Learning Outcome : Recognizes terms and figures.

Q : Following figures are all examples of :

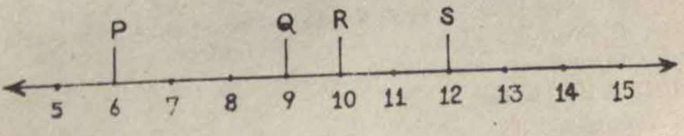


A. Squares B. Quadrilaterals C. Parellelograms D. Trapeziums ( )

8. *Objective* : Skill *Topic* : Number line.

Learning Outcome : Reads graphs.

Q : What is the measure of segment PR + segment QS ?

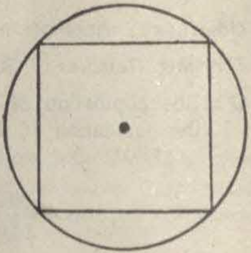


A 6 B. 7 C. 8 D. 5 .. ( )

9. *Objective* : Application *Topic* : Cyclic figures

Learning Outcome : Analyses a problem.

Q : If **R** is the radius of the circle in the adjoining figure, then the side of the square inscribed in the circle is :



A. 2R B. R C. $2R\sqrt{3}$ 3. $R\sqrt{2}$ .. ( )

10. *Objective* : Application *Topic* : Ratio and Proportion

Learning Outcome : Analyses a problem.

Q : Two persons M and N run a concern with capitals of Rs. 200 and Rs. 300 respectively. If M's money has been invested for 3 months and N's for 2 months, the profits will be divided in the ratio of

A. 2:3 B. 1:1 C. 3:2 D. 4:3 .. ( )

11. *Objective* : Application *Topic* : Ratio and Proportion

Learning Outcome : Carries out process correctly.

Q : In a mixture of 100 gallons, milk and water are in the ratio of 7:3. How much milk should be added to bring the ratio to 3:1 ?

A. 10 gallons B. 20 gallons

C. 5 gallons D. 15 gallons .. ( )

12. *Objective* : Understanding *Topic* : Equations and Problems

Learning Outcome : Substitutes necessary symbols and formulae.

Q : To solve the equations $x\text{-}y = 1$ and $3x - 5 = 5$, what value of y should be substituted in the second equation ?

A. $x - 1$ B. $1 - x$ C. $3x - 5$ D. $3x + 5$ .. ( )

13. *Objective* : Application *Topic* : Triangles

Learning Outcome : Reasons deductively.

Q : In a $\triangle$ PQR, O is a point inside such that triangles QOR, ROP and POQ are equal in area. O is the point of intersection of the

A. perpendicular bisectors of the sides.

B. angular bisectors.

C. Medians

D. perpendiculars from the verticles to the opposite sides .. ( )

14. *Objective* : Application *Topic* : Triangles

Learning Outcome : Supports his conclusions by reasons

Q : In an isosceles $\triangle XYZ$, $XY = XZ$, XY is produced beyond X to S such that $XS = XY$. Which of the following should be applied to prove that SZ is perpendicular YZ at Z ?

A. If the sum of two angles of a $\triangle$ is 90, the third angle is a right angle.

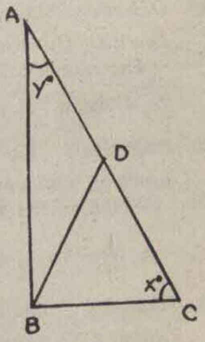
B. If the sum of the angles of a triangle is equal to the third, the third angle is a right angle.

C. If the adjacent angles made by one st. line standing on another st. line are equal, then each is a right angle.

D. If an angle is half of a straight angle, it is a right angle. ( )

15. *Objective* : Application *Topic* : Triangles

Learning Outcome : Establishes relationship



Q : In the given figure

$$DA = DB = DC$$

which of the following is true ?

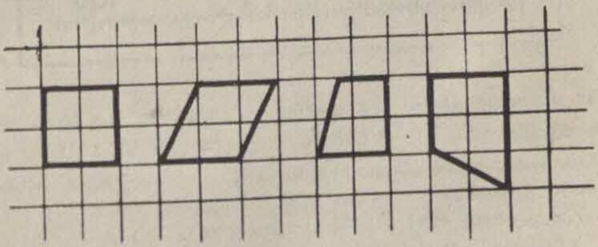
A. $y - x = 0$ B. $y = 2x$

C. $2x + 2y = 180°$ D. $x - y = 90$ ( )

16. *Objective* : Skill *Topic* : Parellelograms

Learning Outcome : Interprets graphs

How many of these figures are parellelograms ?



A. 4 B. 1 C. 0 D. 2 .. ( )

17. *Objective* : Understanding *Topic* : Simple Equations

Q : A student solves the equation

$\frac{x}{5} - 2x + 3 = \frac{2}{3}$ in the following manner. Indicate the step in which he went wrong.

A. $\frac{x}{5} + 3 = \frac{2x}{3} + \frac{2}{3}$ B. $\frac{x}{5} + 3 = \frac{2x+2}{3}$

C. $3x + 45 = 10x$ D. $45 = 7x$ .. ( )

18. *Objective* : **Understanding** *Topic* : Symbolic expressions

Learning Outcome : Translates verbal statements into symbolic statements

Q : A man purchases m tables at n rupees each and he finds his purse short of p rupees to pay the cost of furnitures. Then he must have possessed rupees :

A. $m + n - p$ B. $m + n + p$

C. $m - n - p$ D. $m \times n + p$ .. ( )

19. *Objective* : Knowledge *Topic* : Fractions

Learning Outcome : The pupil recognises terms

The numerical 7 in is called

A. division B. denominator C. quotient D. numerator .. ( )

20. *Objective* : Knowledge *Topic* : Fractions

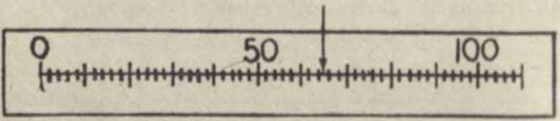
Learning Outcome : Recalls processes

Which is the largest fraction

A. $\frac{1}{3}$ B. $\frac{1}{6}$ C. $\frac{1}{5}$ D. $\frac{1}{4}$ ... ... ( )

21. *Objective* : Skill *Topic* : Graphs

Learning Outcome : Reads graph correctly



On the scale above, the reading indicated by the arrow is between

A. 62 and 64 B 68 and 70 C. 64 and 66 D. 62 and 63

22. *Objective* : Skill computational skill *Topic* : Multiplication

Learning Outcome : Does oral calculations quickly

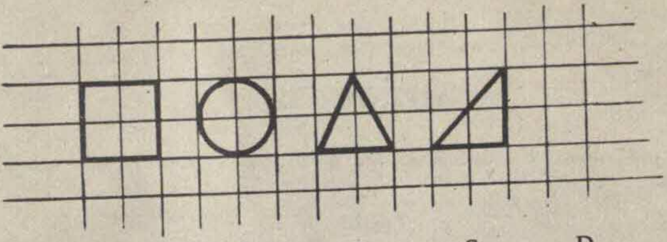
The product 4991 × 191 is closest to

A. 1,000,000 B. 100,000, C. 10,000, D. 10,000,000 .. ( )

23. *Objective* : Skill *Topic* : Area

Learning Outcome : Interprets graphs

Q : Which figure has the greatest area ?



A B C D .. ( )

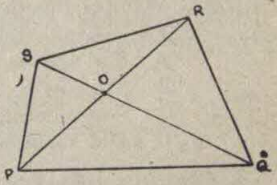
24. *Objectives* : Application *Topic* : Concyclic figures

Learning Outcome : Analyses a problem

Q : If in figure, the points PQRS are concyclic, then ∠ ORQ =

A. ∠ ROQ B. ∠ OPQ C. ∠ RSO

D. None of the three.



Answers

1	D	2	B	3	B	4	C	5	B	6	D	7	B	8	B
9	D	10	B	11	C	12	A	13	C	14	B	15	C	16	D
17	C	18	A	19	D	20	A	21	C	22	A	23	A	24	D

APPENDIX II

A few objective based items are given below from various school subjects.

English

1. *Objective* : Knowledge of the elements of language

   Learning Outcome : Recognizes the correct form

   Q : She was too excited to speak.

   The complex form of this statement is

   A. She was so excited as to speak
   B. She was so excited that she could not speak
   C. Her excitement did not allow her to speak
   D. Her excitement was too great for words.

2. *Objective* : Reading with comprehension

   Learning Outcome : Graphs the meanings of words

   Q : Some dogs love the water. When they see a pond they love to ——— in and ——— around

   A : eat, bark, jump, drink
   B : climb, swim, run, swing

3. *Objective* : Writing simple English correctly

   Learning Outcome : Uses varied sentence patterns

   Q : When the main clause in a sentence is 'He said that', the subordinate clause should be

   A. he will come B. he would come
   C. he will be coming D. he should come

4. *Objjective* : Writing simple English correctly

   Learning Outcome : Uses correct forms

   Q : The school bell (rung, rang, rings) and the recess ended. Sunil and Anita stopped talking about (there, their's their) plan for the holidays.

Geography

1. *Objective* : Knowledge *Topic* : Natural Regions

   Learning Outcome : Recognizes facts

   Q : Which of the countries does not have either a cold or a hot desert area ?

   A. U.S.A. B. China C. Brazil D. Australia

2. *Objective* : Application *Topic* : North America

   Learning Outcome : Predicts consequences

Q : In the event of the closure of Panama Canal, the foreign trade of U.S.A. will suffer most in respect of :

A. Cotton B. Iron-Ore C. Wool D. Coal .. ( )

3. *Objective* : Knowledge *Topic* : Physical Geography

Learning Outcome : Recalls concepts

Q : The latitude of a place is its distance from

A. The poles B. The equator—north or south

C. Greenwich D. The sea .. ( )

4. *Objective* : Understanding *Topic* : Communications

Learning Outcome : Identifies cause and effect relationship

Q : Which of the following conditions hinders the construction of a railway line between Bombay and Goa on the western coast of India ?

A. Hilly area B. Non-availability of cheap labour

C. Heavy rains D. Presence of a net-work of roads .. ( )

Civics

1. *Objective* : Knowledge *Topic* : Economic Movements

Learning Outcome : Recognizes principles

Q : In a co-operative society

A. Membership is compulsory B. Responsibility is individual

C. Government wields all power D. Members take functional decisions .. ( )

2. *Objective* : Application *Topic* : UNO

Learning Outcome : Predicts

Q : Two of the member countries of UNO do not desire to solve their border dispute through the UNO but are preparing themselves militarily. The international body which is competent to deal with the problem is

A. International Court of Justice B. Disarmament Commission

C. Security Council D. General Assembly .. ( )

3. *Objective* : Application *Topic* : Indian Constitution

Learning Outcome : Predicts the outcome of the situation

Q : A bill proposing to amend the constitution is sent for assent to the President of India after having been passed by both the houses of the Parliament. If the President disagrees with the contents of the bill he could —

A. return the bill for reconsideration B. outright reject the bill

C. dissolve the Parliament D. send the bill to the PM for his opinion .. ( )

4. *Objective* : Understanding *Topic* : Indian Constitution
Learning Outcome : Interprets the functions
Q : The effective control over central finance rests with
A. The Rajya Sabha B. The Finance Minister
C. The Lok Sabha D. The President of India .. ( )

Science

1. *Objective* : Understanding *Topic* : Archimedes Principle
Learning Outcome : Selects appropriate illustrations
Q : Which one of the following statements is an illustration of the principle of Archimedes ?
A. A country boat floats on water
B. A piece of iron sinks into water when dropped into it
C. A stone seems to weigh less in water than in air
D. Icebergs float on water .. ( )

2. *Objective* : Application *Topic* : Physical and Chemical Changes
Learning Outcome : Locates principle in a given situation
Q : Which one of the following is an example of a chemical change ?
A. Sugar is dissolved in water B. Ice-cream is prepared from milk
C. A wire is heated D. Curd is prepared from milk ( )

3. *Objective* : Knowledge *Topic* : Measuring Volume
Learning Outcome : Recognises
Q : Which of the following instruments would be more suitable for measuring 32·6 cc. of a liquid ?
A. A measuring jar B. A measuring cylinder
C. A pipette D. A burette .. ( )

4. *Objective* : Understanding *Topic* : Reptiles
Learning Outcome : Discriminates
Q : A snake is a reptile while an earth-worm is not because it has no
A. Poisonous teeth and tongue B. Hood and eyes
C. Scales and backbone D. Eyes and ears ( )

5. *Objective* : Application *Topic* : Reflection
Learning Outcome : Relates basic principles in solving problems
Q : A clock hung on a wall has marks instead of numbers on its dial. On the opposite wall, there is a mirror and the image of the clock in the mirror if read, indicates the time as 7-20. Which of the following is the correct time ?
A. 5-20 B. 4-40 C. 5-40 D. 4-20 .. ( )

Answers :

English	1 – B. 2 – A—jump. B – swim. 3 – B. 4 – rang. their
Geography	1 – C. 2 – A. 3 – B. 4 – A
Civics	1 – D. 2 – C. 3 – A. 4 – C
Science	1 – C. 2 – D. 3 – D. 4 – C. 5 B

* * *

INDEX

A

Accomplishment quotient 168
Achievement tests 196
  diagnostic 196
  essay 192–189
  objective 190–192
  oral 192
  performance 193
  standardized 196
  survey 196
  uses 199
Adequacy 151, 159
Adjustment inventories 239
Administration of tests 147
Age norms 165
Alternate response test 209
Anecdotal record 17, 234, 235
Application objectives 50, 56, 69
Attitude 41, 47, 242

B

Batteries 196
Behaviour changes 26, 32, 51–54
Bias 232
Bimodal curve 300, 356
Bloom Benjamin S. 38–42
Blue print 121, 126, 136

C

Central Tendency
  mean 288
  median 294
  mode 300
Checklist 17, 236
Class-interval 272
Class-room tests (see also essay type, objective type, unit test etc.) 179
Column diagram 277
Completion test 206–209
  maxims 207
  merits and limitations 207
  uses 206
Correction for guessing 146
Correlation coefficient 331
  computation 333
  interpretation 336
  meaning 331
  uses 338
Cumulative record 17, 237

D

Derived scores 163
  (see also Norms, Standard scores)
Design of a Test 119, 135
Deviation 311–325
  average 316
  quartile 313
  standard 319
Deviation IQ 170, 361
Diagnostic tests 196
Difficulty value 133, 137, 140–144
Directions for tests 145
Discriminating power 140–144, 161
Distractors 142
Distribution
  frequency 271–284
  normal 346
  skewed 305, 352

E

Editing test items 122
Educational age 168
Educational objective 27, 29
Educational quotient 168
Essay tests 183–189, 195
  limitations 183
  merits 183
  suggestions for improving 186
  suggestions for scoring 189
Evaluation
  approach 12
  implications 18
  meaning 9
  principles 11
  teaching-learning 13
  tools and techniques 15, 18
  uses 20
Examinations
  defects 180

F

Factors affecting
  reliability 156
  validity 152
Feed back 14
Frequency distribution 271–284
  graphic representation 277

    range 272
    tabulation 273
Frequency Polygon
    construction 279
    smoothing 281

G

Grade 172–175
Grade norms 166
'Guess Who' technique 250
Guessing 192

H

Halo effect 154, 185
Histogram 277
Horns effect 185

I

Instructional objectives 50
Intelligence Quotient 167
Interest inventories 18, 241
Internal Assessment
    areas 256
    importance 265
    plan 256
    validity 258
    weightage 257
Interpretation 161–170
Interpretive exercise 222–228
Interquartile range 313
Interview 17, 238
Inventories 17, 239
Item analysis 124, 138–144
    discriminating power 140–144
    difficulty value 133, 140–144
    validity 151–155
Items
    alternate response 209
    completion 206
    interpretive 222
    matching 213
    multiple-choice 217
    writing of 136

K

Knowledge objective 50, 78
Kuder Reference Record 241
Kurtosis 354

L

Learning Experiences, 13, 15, 27, 84–102
Learning Outcomes 13, 28, 30, 54–75, 78–80
Length of tests 156
Lesson evaluation 259

M

Marking and reporting 172
Marking scheme 123, 132
Marks 172
Matching item 213–216
    maxims for writing 213
    merits and limitations 213
Mean
    assumed 291
    computation of 288
    uses 306
Measurement 9, 161
Median
    computation of 295
    meaning of 294
Mental age 167
Mode 300–302
Multiple choice test 217–221
    maxims for writing 220
    merits and limitations 219

N

Normal curve 346–351
    application 349
    area 349
    properties 348
Norms 147, 165–170
    age 165
    grade 166
    percentiles 169, 361
    quotients 167
    standard scores 170, 358
    types 165

O

Objectives
    classification 22, 35, 50
    criteria for writing 33–35
    defining 13
    instructional 50, 190
    meaning 23, 27
    objective-based 15, 16, 190
    specifications 26, 30, 33, 51–75
    statement 26

taxonomy 37
Objective type tests 190–192, 195
completion 206
interpretive 222
limitations 191
matching 213
merits 190
multiple choice 217
short answer 206
suggestions for writing 203
true-false 209
Objectivity 151, 156, 159
Observational technique 17, 229, 237
anecdotal records 17, 234
checklists 17, 236
cumulative records 17, 237
rating scales 17, 229
Oral examination 192, 195

P
Percentiles rank 170, 361
Percentiles 169, 171, 362
Performance test 193, 195
Personality inventory 17, 239
Product 28
Product evaluation 194
Projective technique 18, 243

Q
Qualitative 10, 15, 21, 229–253
Quantitative 10, 15, 21, 178–201, 203–228
Quartile deviation 313
Quartiles 315
Questionnaires 18, 239
(see also inventories)
Quotients 167–169
AQ 168
EQ 168
IQ 167

R
Range 312
Rank difference 333
Rank order 358
Rating scales 17, 229
common errors 232
reliability and validity 233
types of 230
Raw scores 163
Reading comprehension 133
Reliability 151, 155–158, 185
factors influencing 156
meaning of 155
methods of 155
relation between 157
Reliability coefficient 155
Reporting 172–175
Rorschach 244

S
Scales
attitude 242
rating 229–233
Scores
age equivalent 165
derived 163
grade equivalent 166
interpretation of 161–173, 357
percentile 169, 362
raw 163, 357
standard 170, 358
Scoring
correction for guessing 146
ease of 160
essay tests 189
key 123, 132, 146
Self-report technique
attitude scales 17, 242
interest inventories 18, 241
interview 239
questionnaire 17, 239
personality inventories 17, 239
Semi-interquartile range 313
Short answer test
characteristics and limitations 207
suggestions 207
uses 206
Skewness 352
Skills 50
Sociogram 250
Sociometric technique 249
Specific determiners 211
Specifications 51
Speed tests 137, 145
Split half method 156
Standard deviation
characteristics of 324

  computation 319
  interpretation 321
  merits and limitations 324
  uses 325
Standard scores 170, 358
  deviating IQ 170, 361
  normal distribution 366
  T-scores 361
  z-scores 360
Standardization of a test
  blue-print 136
  design 135
  item analysis 138–144
  norms 147
  preliminary and final form 136
  scoring key 146–147
  validity and reliability 148
Survey tests 196

T

Taxonomy of Educational Objectives
  affective 41–43, 101
  cognitive 39–40, 101
  meaning 37–39
  psychomotor 43, 101
  tripartite structure 45
  uses of 45
TAT 245
Teacher education 77–83, 99
Terman Lewis M. 167
Test
  administration 147
  blue-print 121
  difficulty 137
  diagnostic 196
  essay 183
  objective 190
  oral 192
  performance 193
  survey 196
  standardized 196
  uses of 199
Test items
  completion 206
  criteria 122
  interpretive 222
  matching 213
  multiple choice 217
  true false 209
Tools 15, 16, 18, 179
Triangle of education 13, 263
T-scores 171, 361

U

Understanding 50
Ungrouped scores 270
Unit Plan
  aspects 105
  content analysis 115
  evidence 108–115
  meaning 164
  learning experiences 108–115
  objectives 108–115
Unit Test
  blue-print 121, 126
  design 119
  editing 122
  item analysis 124
  marking scheme 123, 132
  question papers 125–131
  questionwise analysis 124, 133
  scoring key 132
Usability 151, 160, 186

V

Validity 148, 151–154, 183, 233
  content validity 124
  factors influencing 151
  meaning of 151
  relation to reliability 157
Variability
  average 316
  quartile 313
  range 312
  standard 319

W

Weightage 120, 125, 136

Y

Year's plan 116

Z

z-scores 171, 360
zero point 162

* * *

BIBLIOGRAPHY

Bartz, Albert E.	Educational Measurement
Bean Kenreth L.	Construction of Educational and Personnel Tests
Blair, Gleen, Myers	Diagnostic and Remedial Teaching
Bloom Benjamin S.	Taxonomy of Educational Objectives [Cognitive Domain]
Ebel R. L.	Measuring Educational Achievement
Furst Edward J.	Constructing Evaluation Instruments
Garrett Henry E.	Statistics in Education and Psychology
Green, Jorgensen and Gerberich	Measurement and Evaluation in the Secondary Schools
Green J. A.	Teacher Made Tests
Gronlund Norman E.	Measurement and Evaluation in Teaching
Guilford J. P.	Fundamental Statistics in Psychology and Education
Jordan A. M.	Measurement in Education : an Introduction
Lindeman R. H.	Educational Measurement
Lindvall C. M.	Measuring Pupil Achievement and Aptitude
Manuel H. T.	Elementary Statistics for Teachers
Noll Victor H.	Introduction to Educational Measurement
Palardy J. Michael	Elementary School Curriculum
Remmers, Gage and Rummel	A Practical Introduction to Measurement and Evaluation
Risk Thomas M.	Principles and Practices. of Teaching in Secondary Schools
Stanley J. C.	Measurement in To-day's Schools
Thorndike and Hagen	Measurement and Evaluation in Psychology and Education
Travers	Educational Measurement
Wrightstone, Justman and Robbins	Evaluation in Modern Education

By the Same Author

(1) Fundamentals of Experimental Psychology (3rd revised Edition, 1978) Moghe Prakashan, Kolhapur (Maharashtra)

(2) Measurement in Education and Psychology, Moghe Prakashan, Kolhapur (Maharashtra)

(3) Psychological Foundations of Education (2nd revised Edition, 1985) Macmillan India, Bombay 400 010

(4) Elementary Statistics in Education and Psychology (1979) Sheth Publishers, Bombay 400 026

(5) Mental Ability Tests (NTS Examination) (1985) Anamol Prakashan, Pune 411 002

(६) शैक्षणिक व प्रायोगिक मानसशास्त्र (तिसरी सुधारित आवृत्ती १९८१) मोघे प्रकाशन, कोल्हापूर

(७) शैक्षणिक मूल्यमापन व संख्याशास्त्र (दुसरी आवृत्ती १९८५) श्रीविद्या प्रकाशन, पुणे ४११ ०३०

(८) शैक्षणिक मूल्यमापन (पाचवी आवृत्ती १९७६) श्रीविद्या प्रकाशन, पुणे ३०

(९) बुद्धिमापन कसोट्यांची तोंडओळख (भाग १ : चौदावे पुनर्मुद्रण १९८४) श्रीविद्या प्रकाशन, पुणे ३०

(१०) बुद्धिमापन कसोट्यांची तोंडओळख (भाग २ : सुधारित आवृत्ती, चौथे पुनर्मुद्रण १९८३) श्रीविद्या प्रकाशन, पुणे ३०

(११) बुद्धिमत्ता विकास व चाचणी (प्रारंभिक व भाग १ : १९८१) मयूर प्रकाशन, पुणे

(१२) मानसिक क्षमता कसोटी (१९८२) अनमोल प्रकाशन, पुणे २

(१३) अनमोल मिडलस्कूल स्कॉलरशिप परीक्षा (१९८३)

(१४) अनमोल हायस्कूल स्कॉलरशिप परीक्षा (१९८४)

(१५) राष्ट्रीय प्रज्ञाशोध परीक्षा : मानसिक क्षमता कसोटी (१९८५) अनमोल प्रकाशन, पुणे ४११ ०३०

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

* * *

Rs. 50-00

371.26
DAN

F 13

SHRIVIDYA PRAKASHAN
250 Shaniwar, Ashtabhuja Marg, Poona 30